THALES

12om Limited Objective Experiment #2
Final Results Summary and Recommendations

Prepared by:

Thales Canada, Defence and Security
1405 boul. du Parc-Technologique
Québec, QC G1P 4P5

C3 Human Factors Consulting Inc.
1405, boul. Du Parc-Technologique
Québec, QC, G1P 4P5

For:

CSA: Natalia Derbentseva
Claire Lalancette and Michel Lizotte
Defence Research and Development Canada – Toronto
1133 Sheppard Ave. W
Toronto, Ontario M3K 2C9

31 March 2014

Thales Document Control Number (DCN): 2014C.007-REP-01-AT7 Rev. 01

DRDC-RDDC-2014-C98

CONTRACT REPORT

Authors: Jean-François Gagnon
Daniel Lafond
Marie-Ève St-Louis
Sébastien Tremblay

Scientific Authority: Natalia Derbentseva

ABSTRACT

12om is a Defence Research and Development Canada (DRDC) project that seeks to develop a methodology (i.e., process and components) to improve the understanding of a complex situation by a multidisciplinary, government-wide team. The 12om methodology (see report 2014C.005-REP-07-AT5) was developed through a combination of empirical and theoretical work. The objective of the current Limited objective experiment (LOE) was to conduct a team-centred evaluation in the context of the current Canadian Forces Operational Planning Process (CF OPP) for a Joint Plans (J5) Whole of Government (WoG) planning team and assess the methodology with regards to the functions it aims to support. In the tested version, 12om comprised seven components: (1) WoG OPP handbook, (2) team building and handover procedure, (3) interactive common glossary, (4) conceptual diagrams (individual and collaborative), (5) cross-impact-method, (6) OP Design tool, and (7) integrated mission analysis briefing template. An integrated planning team of five members took part in this study. The approach adopted to guide the current evaluation was based on Cognitive Systems Engineering (CSE; Rasmussen, Pejtersen, & Goodstein, 1994; Woods & Roth, 1988). CSE assesses, from a holistic perspective, cognitive work in complex domains through human in-the-loop testing. Various measures were taken during the experiment. Among others, these measures assessed the usefulness, usability, costs/benefits ratio, and time and effort required to use and train each of the 12om components. They also assessed the level of support of the 12om methodology as a whole with regards to three critical dimensions: integration of perspectives, collaboration, and common understanding. Other measures aimed to assess mediating factors on team performance such as workload (assessed with the NASA-TLX), team dynamics (assessed through a social network analysis of the team's communication during the experiment) and transactive memory systems. A task-to-tool mapping allowed conducting a functional gap analysis. Finally, measures of effectiveness were collected through a comprehensive evaluation of the planning team's output and process by subject matter experts. Some of these measures were integrated within multicriteria hierarchical models in order to provide (1) individual assessments of the components of the methodology and (2) an overall assessment of the method with regards to its main objectives. The use of the multicriteria models also allowed performing a sensitivity analysis providing diagnostics on the key leverage points for improving the methodology or its components. In terms of 12om methodology overall and component assessments, results were highly positive, demonstrating the potential of the approach for supporting collaborative understanding of complex situations in the context of multidisciplinary teamwork. The functional gap analysis revealed that the 12om methodology covered a broad spectrum of the OPP tasks. It also demonstrated that impact on taskwork varies across components and that all components are not suited for supporting all the OPP tasks. These results are consistent with a toolbox approach to the use of the 12om methodology components, where the use of specific components is handpicked given the characteristics of the task.

EXECUTIVE SUMMARY

The present document outlines the activities and associated observations made during the 12om LOE #2, held at DRDC-Valcartier, on January 13th - 17th 2014. The aim of the LOE was to collect feedback on the 12om methodology as a whole and on its individual components. The 12om methodology seeks to increase performance of inter-agency teams through the support of three dimensions: the collaboration processes, integration of different perspectives, and common understanding. The methodology was consequently assessed in relation to these objectives. The individual components were assessed in terms of costs (e.g. effort to use the tools, time necessary to use/train the tools) vs. benefits (e.g., usefulness of the tools with regards to the 12om objectives). Additionally several measures on the team process and products were used to comprehensively characterize the outcomes of the LOE.

Specifically, the 12om methodology was applied to the first phases of the CF OPP cycle in the context of a simulated WoG effort carried out to respond to a polio outbreak in the horn of Africa (i.e., Somaliland, Ethiopia, Djibouti). A WoG team of five members was created for that purpose (three from the CF [J5, J52; J5Ops] and two civilians [J5Dev and J5Gov]). They had to perform a mission analysis as well as an initial course of actions (COA) development. They had to report to the ROC and Commander, which evaluated their efforts in terms of process and output. Four observers also evaluated their performance. Two of the observers were high ranked officers from the CF, and two were civilians with extended experience in this work domain. A constellation of measurement tools (i.e. camera feeds, paper questionnaires and open/directed discussions) allowed for a comprehensive assessment of the 12om methodology.

The main results of the study are summarised as follows:

- 12om methodology was rated very favourably:
  - Its main strength is in supporting the three dimensions of Collaborative Understanding (i.e., integration of perspectives, collaboration support, and common understanding) in a very balanced way;
  - Its support covers all tasks associated with the OPP;
  - Its support covers all collaborative understanding dimensions relevant to other multi-agency and/or planning contexts; and
  - Analyses reveal that future work on 12om methodology should prioritize the improvement of the development of common understanding.
- The WoG team met expectations in terms of effectiveness:
  - All indicators of performance and effectiveness including team dynamics, workload, and situation awareness, were positive.
- All components were well rated in terms of compromise between support and feasibility of implementation.
- Areas for improvement of the 12om methodology component were identified.
- Overall, the LOE demonstrated the high maturity level of the 12om methodology:
  - All ratings were positive (objective and subjective); and

- Canadian Forces have already made an official request to DRDC for the right to use this toolkit in its future interagency planning endeavors.

TABLE OF CONTENTS

TABLE OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

Acronym	Definition
ARP	Applied Research Project
CD	Conceptual Diagrams
CF	Canadian Forces
CFD	Chief of Force Development
CG	Concept Graphs
CIM	Cross-Impact Method
COA	Course of actions
CONOPS	CONcept of OPerationS
CSE	Cognitive Systems Engineering
Dev	Development
DFATD	Department of Foreign Affairs, Trade and Development
DP	Decisive Point
DRDC	Defence Research and Development Canada
FGS	Federal Government of Somalia
Gov	Governance
HREC	Human Research Ethics Committee
IDP	Internally Displaced Persons
KR	Knowledge Representation
LOE	Limited Objective Experiment
LOC	Lines of Communication
MA	Mission Analysis
MARS	Mission Awareness Rating Scale
NASA-TLX	NASA-Task load index
NGO	Non Government Organisation
OP	Operational Planning
OPP	Operational Planning Process
Ops	Operations
ROC	Representative of Canada
SA	Situation Awareness
SME	Subject Matter Experts
TFC	Task Force Commander
TMS	Transactive Memory Systems

WoG	Whole of Government

1 Introduction

The document presents the results of a LOE to assess teamwork and performance aspects of a new methodology for process planning in support of the Applied Research Project (ARP) entitled "12om". The overall purpose of this project is to develop a *methodology* (i.e., a combination of processes and support tools) to improve the understanding of a complex situation by a multidisciplinary team combining experts from different governmental departments. Specifically, 12om methodology seeks to support:

1. The integration of different perspectives;
2. The collaboration process between multidisciplinary team members; and
3. The ability to develop a common understanding.

1.1 Background

Collaborative mission analysis is an inherently difficult enterprise especially in a WoG planning context. The 12om project seeks to identify and address these challenges through the following technical objectives:

1. Identify support requirements in LOE #1, which will provide input into the design and development of support tools and processes to improve team collaboration within this type of context; and
2. To pilot the developed methodology in LOE #2. The present document reports the results of LOE #2 specifically.

The collaboration context selected is a Joint Plans (J5) integrated (WoG) planning group. Furthermore, the project focus is on the mission analysis (orientation phase of the operational planning process) and COA development phases of the CF OPP.

1.2 Operational Planning Process

The CF OPP is comprised of five main stages (see Figure 1):

- The *Initiation* stage results in the activation of the planning staff and the commander's guidelines about the kind of planning process to achieve;
- The *Orientation* stage results in the development of the commander's planning guidance. At this stage, the commander orients his/her staff towards the determination of the nature of the problem and the confirmation of the results to be achieved;
- The *COA Development* stage results in the production of the CONOPS (CONcept of OPerationS) that identifies the commander's line of action in order to accomplish his/her mission. It presents the COA that will be implemented;
- The *Plan Development* stage results in a set of orders based on the commander's decision to provide subordinate and supporting units with all of the necessary information to initiate the planning or the execution of operations; and

- The *Plan Review* stage results in a regular review of the plan to evaluate its viability. The review period of the plan depends on the evolution of the situation, the type of operation and the environment.

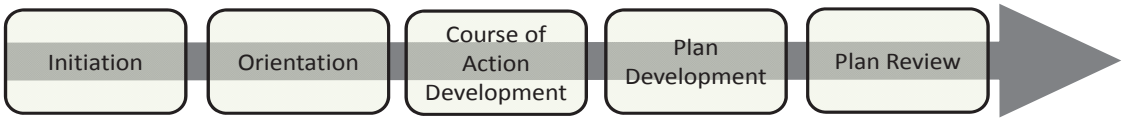


Figure 1: Five phases of the Canadian Forces Operational Planning Process

Figure 1 also illustrates an ideal process in which the whole OPP is performed flawlessly in terms of the quality and timeliness all five components of the process (i.e., a straight arrow). However, the inherent complexity of planning within a WoG context makes accomplishing this ideal process extremely unlikely. Rather, sub-optimal team dynamics, poor shared awareness and high levels of workload (amongst other team-centred factors) are likely to lead to sub-optimal planning performance (see Figure 2 for an example).

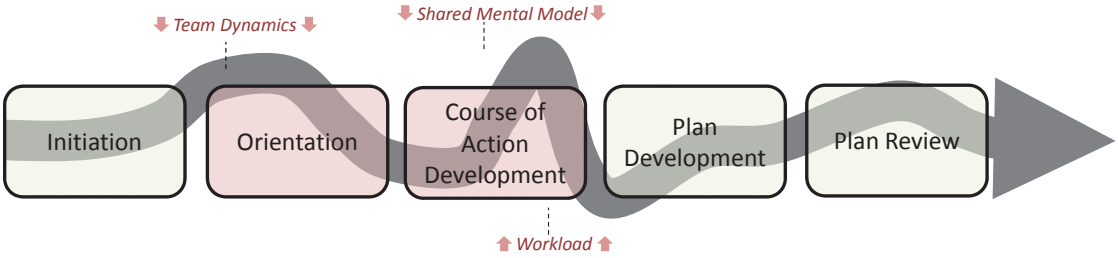


Figure 2: Team-centred factors leading to sub-optimal Orientation and COA Development performance

The overall purpose of the 12om project, therefore, is to develop and evaluate a *methodology* – by which we mean a coordinated set of process improvements and support tools – to improve collaboration in multidisciplinary teams. The context in which the methodology is tested is the CF OPP within a J5 WoG planning group. For example, Figure 3 illustrates a hypothetical scenario within which a combination of process improvements and support tools have been used to address the planning deficiencies identified in Figure 2. In doing so, the overall CF OPP process performance is improved to an acceptable level.

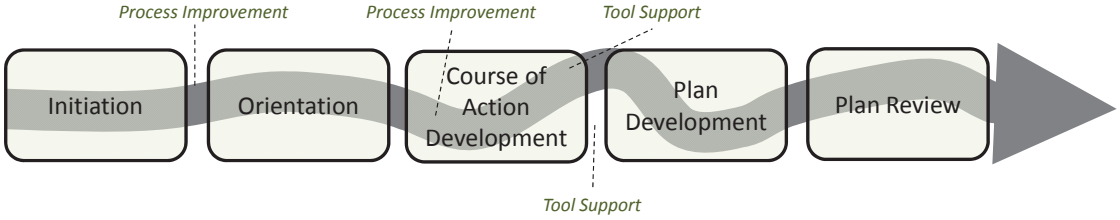


Figure 3: Illustration of how 12om seeks to improve the CF OPP through the development and application of a novel Methodology

1.3 12om Methodology

The 12om methodology is based on a *toolbox* approach (Gigerenzer & Gaissmaier, 2011; Gigerenzer & Selten, 2001; Gigerenzer, Todd & al., 1999). Research in various domains suggest that the strength of a toolbox approach lies on the fact that a given methodology

component is not intrinsically good or bad in enhancing analyst' comprehension, but that it's utility has to be determined in relation to the characteristics and constraints of the task (Dieckmann & Rieskamp, 2007; Rieskamp & Otto, 2006). Such constraints, for instance, are temporal pressure (Karelaia & Hogarth, 2008) or uncertainty (e.g., Marewski & Schooler, 2011). Each component doesn't have to be applied to all contexts as they can be handpicked at will when relevant. For instance, in time-pressured tasks where quick action is required, one might not have the time to develop an integrated representation of the situation with a knowledge representation tool, however, one may benefit from such a representation if the situation does not require quick actions.

Seven components were integrated within the 12om toolbox for LOE #2: (1) WoG OPP handbook, (2) Team building and handover procedure, (3) Interactive common glossary, (4) Collaborative knowledge representation (i.e., individual and collaborative conceptual diagrams), (5) Cross-impact method, (6) Op Design and (7) Mission analysis (MA) briefing template. Selection and integration of these components was based on the assessment by DRDC experts of the critical results from the previous workshops. These components are briefly described and justified below:

- *WoG OPP handbook.* The WoG OPP handbook is a small document designed for to the members of joint civil-military planning endeavors, such as WoG approach initiatives. It comprises a summary of the main phases of the OPP and associated sub-tasks. The purpose of the handbook is to inform joint civil-military teams about what is expected from them. The integration of this component is justified by the observed discrepancies on the understanding of the OPP across team members, especially between civil and military members. The lack of understanding of the OPP by some of the civilian members may be responsible for their relatively smaller contribution to the planning process.

- *Team building and handover procedure*. Team building and handover procedure consists of a set of activities aiming to foster knowledge about the team's goal and objectives, team members' expertise and specific knowledge, and about the process that will be followed by the team. The integration of this component within the 12om methodology is justified in part by the lack of cohesion between members of the team (especially across agencies) which has been outlined by the SMEs during the series of workshops.

- *Interactive common glossary.* The interactive common glossary is a tool that centralizes and defines situation-related terms and acronyms. Its purpose is to foster the use of a common language during the OPP. The integration of this methodology component within 12om is justified by the observed discrepancies between individuals and/or agencies which often lead to miscomprehension and lack of collaboration.

- *Collaborative knowledge representation*. Collaborative knowledge representation consists of a set of tools and techniques that aims to facilitate the integration of

different perspectives on a situation into a single visual representation. Depending on the level of achieved formalism, this representation can be a sole visual description or a graph-based and logically founded description of the situation. The integration of this component within the 12om methodology is justified by the difficulties observed in comprehensively representing complex situations in the context of OPP. Moreover, this component seeks to foster the integration of multiple perspectives, which is also frequently lacking in this context.

- *Cross-impact method.* The cross-impact method is a structured way to analyse a situation and increase awareness of critical interactions. It aims to improve shared awareness of the multiple factors involved, to make salient key divergences in understanding among a team and to minimize tunnel vision by promoting a comprehensive consideration of factors' direct and indirect effects. Moreover, in the suggested implementation, the cross-impact method also promotes option awareness, that is what can be done to improve a particular variable within the system.

- *OP Design tool/process.* OP Design aims to support planners in sequencing decisive points into lines of operations and to identify operational phases with their associated objectives and tasks. It provides the grounds to initiate the thinking required to identify possible branch plans and/or sequel plans where transition conditions are desired. The integration of this component within 12om is justified by the need for initial OP Design during the orientation phase of the OPP and the limited capabilities of the other components to satisfy this need.

- *Integrated MA briefing template.* The integrated MA briefing template intends to help clarify the nature of the output required by each team member and to reduce formatting work for the mission analysis brief. The integration of this component in the 12om toolbox is justified by the fact that a pure military template for MA briefing may not capture aspects relevant to other agencies like the CF and Department of Foreign Affairs, Trade and Development (DFATD).

For further description and justification of the components as well as a deeper description of the 12om process, please refer to report 2014C.005-REP-07-AT5.

1.4 LOE #2 Objectives

The methodology developed as part of the 12om project should maximize the fit across the processes involved in collaborative sense-making, the techniques for knowledge representation, and the tools required to support the expression and sharing of this knowledge across the planning team. In summary, the 12om project seeks to develop a *methodology* (i.e., process and tools) to improve the understanding of a complex situation by a multidisciplinary, government-wide team. Specifically, LOE #2 objectives are to:

1. Assess integration of different perspectives;
2. Assess collaboration process;
3. Assess ability to develop a common understanding;

4. Identify ways to improve individual components of the methodology; and
5. Assess performance.

The study was reviewed and approved by the DRDC Human Research Ethics Committee (HREC) under protocol 2012_003 Amendment 1.

1.5 Document Overview

This document describes the experimental protocol and results pertaining to the evaluation of the 12om methodology for a J5 WoG planning team. Specifically:

- *Section One: Introduction*. This section provides the background, purpose and objectives of the study, together with an overview of the document;
- *Section Two: Evaluation Protocol*. This section describes the approach that guided the development of the protocol, together with a detailed description of the protocol itself (i.e., participant description, schedule, evaluation procedure, and evaluation measures);
- *Section Three*: Evaluation Results. This section provides a detailed description of the results from this study; and
- *Section Four: Conclusions and Recommendations.* This section summarises the results of the study, together with a series of recommendations for the refinement of the methodology and modifications to the evaluation protocol.

2 Evaluation Protocol

This section describes the approach that guided the development of the protocol, together with a description of the protocol itself.

2.1 Overview of Evaluation Approach

The approach adopted to guide the evaluation of the 12om methodology is CSE (Rasmussen, Pejtersen, & Goodstein, 1994; Woods & Roth, 1988) that assesses, from a holistic perspective, cognitive work in complex domains through human in-the-loop testing. CSE seeks to develop knowledge, methods, and tools to guide the design of systems to support human performance in complex settings such as military strategic decision-making. The field of CSE attempts to combine the areas of human-computer interaction, psychology and human factors in order to design and develop socio-technical systems that are human-centred and that take into account both the needs of the user and the constraints imposed by the task (Potter and Rousseau, 2011). CSE is concerned with the effective design of teams, technological tools, or training methods to support cognitive work (Pfautz & Roth, 2006). In line with the CSE framework, we argue that the holistic evaluation of performance in combination with determinants of cognitive work is essential for identifying validated individual and team-centred design requirements and recommendations for tools, procedures and methods. For example, an evaluation approach based solely on performance assessment will not provide any indication of the underlying causes of good (or poor) levels of individual or team performance.

2.1.1 Evaluation Steps

The evaluation protocol described in this section was developed to provide a way of assessing the 12om methodology without relying on a comparison with the data collected during LOE #1. This constraint arose from the changes made to the scenario, team composition, and many other parameters of the LOEs that would render any comparison obsolete. The evaluation protocol was mostly described elsewhere (see report 2014C.007-REP-02-AT7 - Data collection plan and DRDC HREC protocol 2012_003 Amendment 1), but its key elements are reported here as well to ease comprehension of the LOE #2.

2.1.1.1 Step 1: Develop Data Collection Plan

This step comprised the following activities:

(a) Information Gathering. All the information required for producing the experimental documents, observation grids, training and questionnaires was gathered.

(b) Selection of Analyses and Measures. The selection of analyses and measures was guided by the objectives stated above. To be included within the data collection plan, a measure had to inform, at least partly, one or more of the 12om LOE #2 objectives. Measures either inform directly the objectives or serve as input to more complex analyses which in turn will inform one or more of the 12om LOE #2 objectives. The final set of measures and analyses is listed below.

Analyses and their relation with the different measures will be described further in section 2.3 of the current report.

a. Measures
    i. Methodology assessment questionnaire (see Appendix A)
    ii. Component assessment questionnaire (see Appendix B)
    iii. Individual factors identification questionnaire (see Appendix C)
    iv. Team factor assessment questionnaire (see Appendix D)
    v. NASA-Task load index (NASA-TLX) (see Appendix E)
    vi. Transactive memory systems (TMS) (see Appendix F)
    vii. Mission Awareness Rating Scale (MARS) (see Appendix G)
    viii. Task-to-tool mapping (see Appendix H)
    ix. Mission analysis process evaluation (see Appendix I)
    x. Mission analysis product evaluation (see Appendix J)
    xi. Observers' notes
    xii. Recording of communications (and identification of critical timeframes)
        1. Mission analysis
        2. Task-to-tool mapping
        3. Focus group discussions

b. Analyses
    i. 12om Methodology assessment by criterion
    ii. Multi-criteria assessment – Components
    iii. Multi-criteria assessment – Overall
    iv. Performance analysis
    v. Functional gap analysis
    vi. Common understanding
    vii. Team dynamics
    viii. Qualitative assessment

(c) Production of the material. The experimental material (e.g., questionnaires and data collection sheets) was then produced in a "pen and paper" format (as opposed to electronic format) for practical reasons.

(d) Review of material. The material was reviewed by the Technical Authority and revised according to the technical feedback received.

2.1.1.2 Step 2: Conduct Experiment

This step comprised the following activities:

(a) Execution of the Experimental Sessions. Experiment was conducted at DRDC-Valcartier, in a dedicated area without external distractions to ensure observed effects are caused by experimental manipulations. The LOE setup is described below.

(b) Backup and Logging. The data were stored on external drives and the questionnaires were scanned upon completion of the experiment. All raw data were made available to all project personnel through a Sharepoint Site.

2.1.1.3 Step 3: Data Analysis and Reporting

Raw data was acquired through various sources; including video and audio from a webcam, and subjective impressions with questionnaires. These data were either used to inform directly one of these objectives, or included into more complex analyses which in turn informed the objectives.

2.2 Limited Objective Experiment #2

2.2.1 Task

The integrated planning team observed in this study involved five members: three military planners and two civilians representing of DFATD. The task of the integrated team was to produce a mission analysis brief for a Task Force Commander (TFC) and Civilian Representative of Canada (ROC) in a simulated yet realistic Horn of Africa vignette based on a scenario originally developed by the Chief of Force Development (CFD) and enhanced by the Canadian Army. The team was required to use the 12om methodology components during the execution of the task. The collaboration process was evaluated through a battery of measures collected during the LOE.

2.2.2 Participants

2.2.2.1 WoG Team Members

Five individuals were recruited to participate in the exercise, with no restriction in terms of age and gender (standard adult population ranging from 18-65). The five-person planning team in this study consisted of the following roles:

- J5 Lieutenant-Colonel;
- J52 Major;
- J5Ops Major;
- DFATD development subject matter expert; and
- DFATD governance subject matter expert.

The study required that participants possess a precise set of expertise:

- In-depth knowledge of their respective organizations (i.e., planning CF branch, DFATD), their mandates, policies, and objectives;

- Operational deployment experience as representatives of their respective organizations in Canadian or international WoG missions; and
- Experience with military planning process.

In addition to the time requirements for this study (i.e., five days); the above expertise requirements severely constrained the pool of potential participants. These participants were considered on duty for the duration of the study and were compensated for their time by their respective employers.

2.2.2.2 SME Observers

In addition to the WoG team, four additional participants were recruited as follows:

- Senior military officer (at the rank of Colonel) to play the role of the TFC,
- Senior civilian to play the role of the senior civilian ROC,
- Senior civilian to act as an observer during the exercise, and
- Senior military officer (at the rank of Lieutenant Colonel) to act as an observer during the exercise.

The TFC and the ROC jointly evaluated the mission analysis brief that the teams produced at the end of the task. Because the time commitment required for the TFC and the ROC roles were not as demanding, it was possible to recruit individuals for these roles from the respective government departments (active or retired). Participants were fully informed of the details, discomforts and risks associated with the experimental protocol before being asked for their written consent.

2.2.2.3 Experimenters

The experimenter's team was composed of DRDC and Thales/C3 personal.

2.2.3 Experimental Setup

The team was working in a single room with individual desks, a shared display, a whiteboard, and a central table (see Figure 4). The team also had access to a printer located outside the room. Six desktop computers were networked together and given Internet access. Computers were equipped with Microsoft Office, a web browser and were associated with a personal e-mail address. Two cameras and microphones were installed in the room to allow a group of observers to monitor the activities from another room.

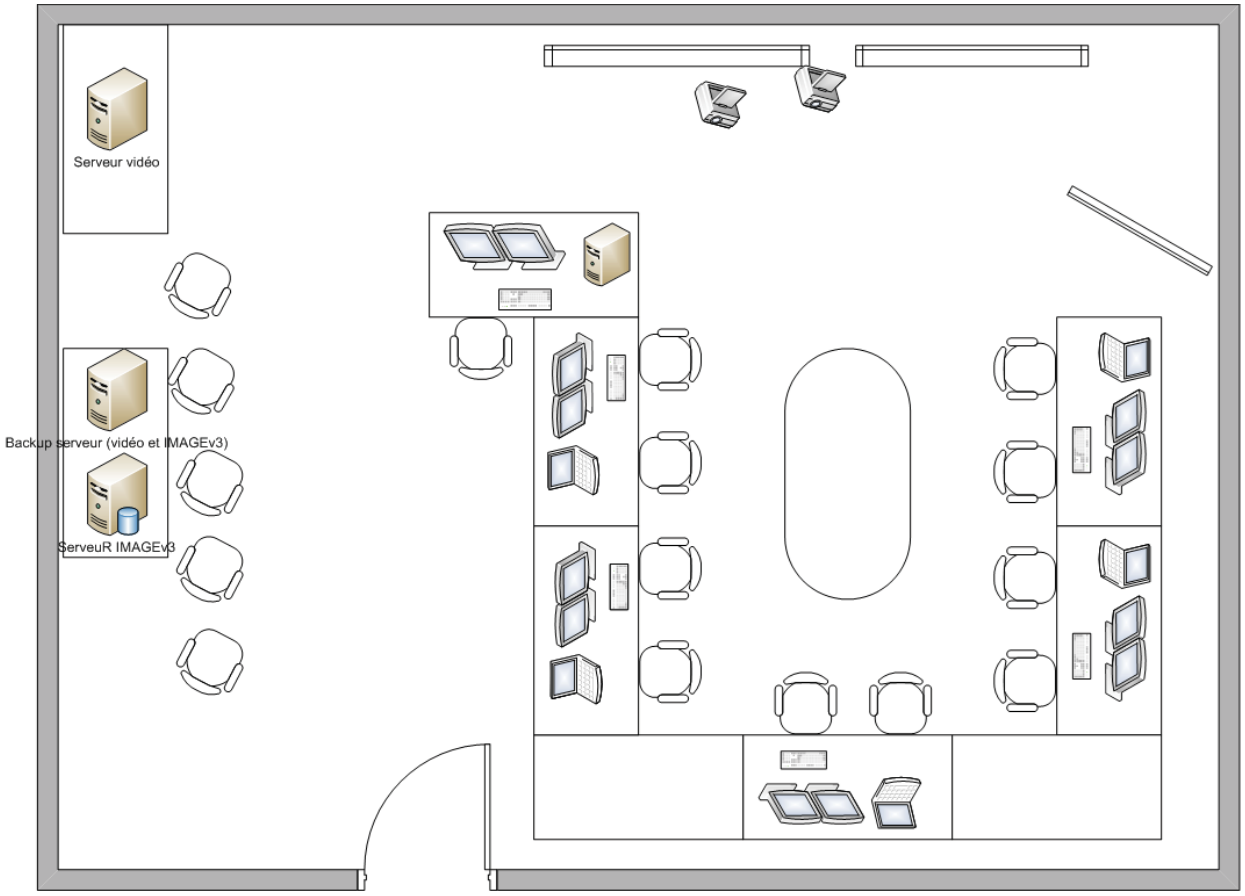


Figure 4: Planning team room layout.

The planning team was observed throughout the duration of the study by a group of observers (i.e., SME observers and experimenters). To reduce any potential distraction caused by their presence, observers were located in a separate room and monitored the group through a live video feed. In addition, the sessions were video recorded to allow for subsequent analysis of the team's interactions. The observers' room layout is depicted in Figure 5.

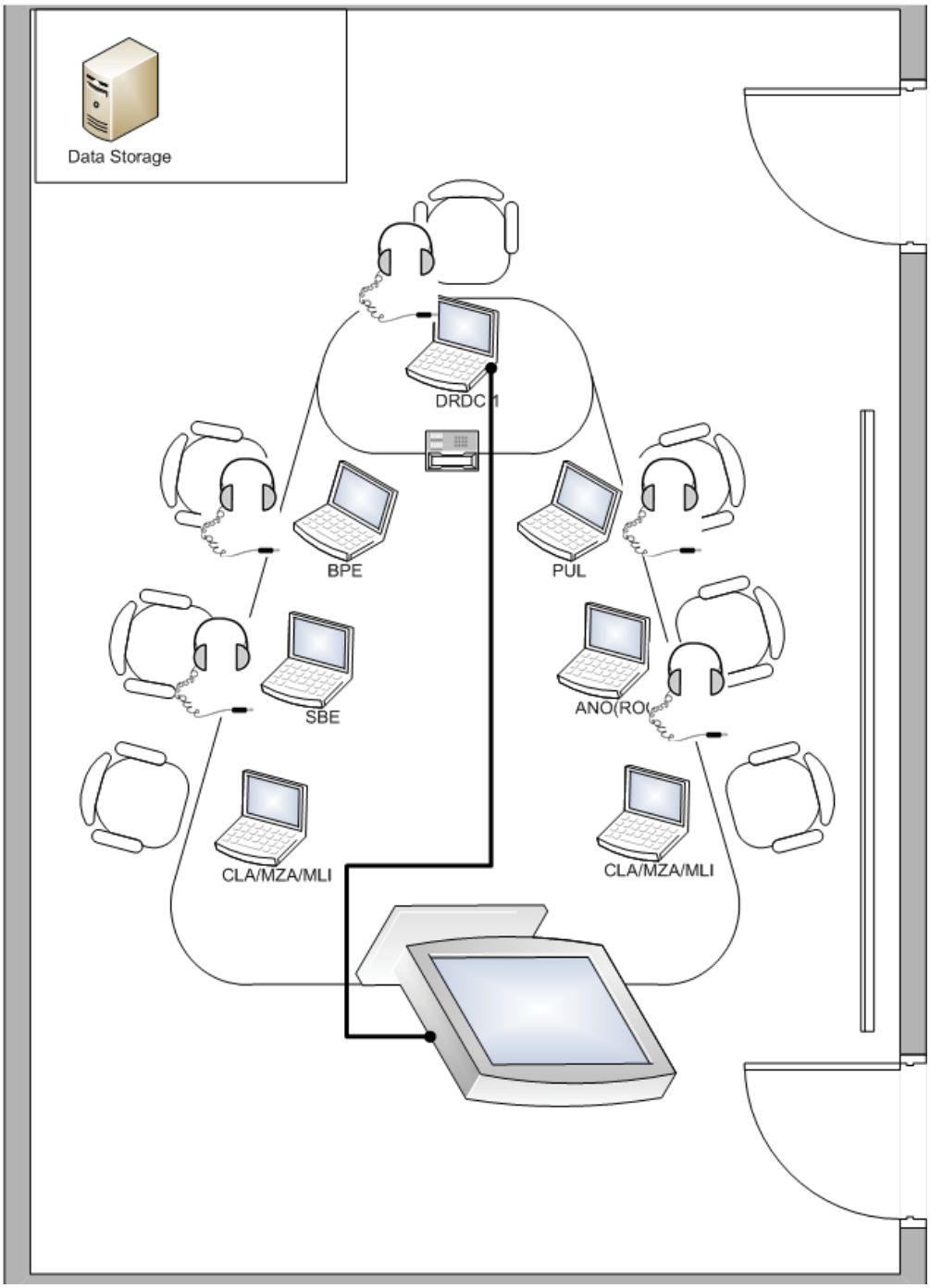


Figure 5: Observers room layout.

2.2.4 Schedule

The experiment lasted four full days spread over five weekdays (January 13th to January 17th). The schedule of the LOE is given below in Table 1. The schedule was mostly respected without any major deviation.

Table 1. Schedule of LOE #2

Day	Activities
Day 1 - 1300-1700	• Introduction • Team building and handover procedure - Part 1 • Glossary – Presentation of the component • Scenario read in
Day 2 - 0800-1700	• Scenario read in (cont'd) • TFC and ROC's Initial guidance LUNCH • Team building and handover procedure - Part 2 • Individual and collaborative knowledge representation – Presentation of the component • Day 2 after action review
Day 3 - 0800-1700	• Collaborative knowledge representation (cont'd) • Mission analysis - undisturbed LUNCH • Mission analysis - undisturbed • OP Design tool – Presentation of the component • Day 3 after action review
Day 4- 0800-1700	• Mission analysis - undisturbed • Mission analysis brief (around 1100) LUNCH • Cross-impact method – Presentation of the component and use • Day 4 after action review • 12om methodology evaluation
Day 5 - 0800-1200	• 12om methodology evaluation

2.3 Measures and Analyses

The experimental approach adopted for this study relies on the assessment of the individual and team behaviour from a holistic point of view, rather than focusing solely on performance. The aim is to gain a comprehensive understanding of the factors associated with performance in this context. Such empirical findings are critical to provide insightful recommendations about how to support the planning process. This section describes analyses carried out, the measures associated with these analyses, and their relevance to 12om second LOE objectives.

2.3.1 12om Methodology Assessment

The core objective of the LOE #2 is to assess the 12om methodology. The main way to collect information with regard to this objective was through the use of questionnaires. All questionnaires are available in annex. The questionnaires were handed out to the WoG team members and SME observers and aimed to assess various dimensions of the 12om methodology. The data collected was used in different ways, described below.

2.3.1.1 Overall 12om Methodology Assessment

One questionnaire aimed at assessing the 12om methodology in terms of the level of support provided to different dimensions it aims to support: (1) collaboration, (2) common understanding, and (3) integration of different perspectives. The following questions were targeted at the 12om methodology as a whole:

- How well does the 12om methodology support the integration of different perspectives?
- How well does the 12om methodology support common understanding?
  and
- How well does the 12om methodology support collaboration?

The 12om Methodology Assessment questionnaire is available in Appendix A.

2.3.1.2 12om Methodology Component Assessment

Although 12om was composed of seven components, the assessment broke down some of them into a series of sub-components. The objective of this breakdown was to collect a more fine grained data on some of the methodology components. The breakdown was as follows:

1. WoG MA Briefing template;
2. OPP handbook;
3. Common glossary;
4. Team building procedure;
5. Knowledge representation (IMAGEv3):
   a. Creation of the common vocabulary;
   b. Creating conceptual diagrams individually;

c. Creating conceptual diagrams collaboratively;
d. Creating views using filters;
e. Querying conceptual diagrams using filters; and
f. Sharing conceptual diagrams.

6. Op Design:
    a. Tool; and
    b. Process.
7. Cross-Impact method.

Individual questionnaires were created to assess each specific components (or sub-components) of the 12om methodology. Overall, 13 components and sub-components were assessed by questionnaires. The goal of these questionnaires was to collect directly the subjective impressions of the team members with regards four dimensions of each component: (a) usefulness, (b) time/Effort required to use and train, (c) usability, and (d) cost/benefits.

(a) Usefulness of the component

This concept is represented by the average ratings (ranging from 0-10) of the participants at the following questionnaire items:

a. Usefulness for supporting the activities and products of the OPP;
b. Usefulness for supporting the collaboration process (in general);
c. Usefulness for supporting the integration of different perspectives (in general); and
d. Usefulness for supporting the development of common understanding (in general).

(b) Time/Effort required to use and train the component

This concept is represented by the average ratings (ranging from 0-10) of the participants at the following questionnaire items:

a. How much time is required (relative to other tools that may be used in this context) for the team to use the COMPONENT?
b. How much effort is required (relative to other tools that may be used in this context) for the team to use the COMPONENT?
c. How much time is required (relative to other tools that may be used in this context) for training on the COMPONENT?

and

d. How much effort is required (relative to other tools that may be used in this context) for training on the COMPONENT?

(c) Usability of the component

This concept is represented by the average rating (ranging from 0-10) of the participants at the following questionnaire item:

a. What is the usability (relative to other tools that may be used in this context) of the COMPONENT?

(d) Cost/benefits function of the component

This concept is represented by the average response for the following questionnaire item:

a. Consider the costs and benefits associated with using the COMPONENT. How much do benefits outweigh the costs?

The mean ratings of the participants at each of these dimensions will be reported by component and sub-component. An example of a component assessment questionnaire is shown in Appendix B.

2.3.1.3 Multi-criteria Assessment

Although the questionnaire items can be very informative on the components/sub-components of the 12om methodology, it is unclear how each item relates to each other and therefore it can be difficult to generate a holistic assessment of either the overall 12om methodology or of a particular component. Some of the items covered above are clearly about assessing the costs of using 12om methodology, for instance in terms of training time and effort, whereas other items focus on assessing the benefits of using the tools, for instance in terms of "support to collaboration". The aforementioned analyses will provide for an overall description of the different ratings for all the 12om methodology components, but not an integrated assessment. The multi-criteria assessment described below precisely aims to integrate all these criteria into a single non-linear utility model.

The MYRIAD tool (Labreuche & Le Huédé, 2005) is used to perform a multi-criteria analysis of the 12om methodology and its individual components. The analysis involves specifying a preference model that captures the design team's priorities using the versatile Choquet integral to combine heterogeneous criteria. This approach can be particularly helpful to create a meaningful synthesis of the key measurement collected during the LOE. The formal framework underlying MYRIAD allows defining how different criteria interact (or not) to produce an overall degree of satisfaction. Criteria can be additive (i.e., satisfaction corresponds to the sum of values), complementary (i.e., satisfaction correspond to the smallest value), or substitutable (i.e., satisfaction corresponds to the highest value).

Overall 12om Methodology Multi-criteria Assessment

Figure 6 shows the preference model used for the general methodology assessment. The variable hierarchy was designed to capture the two main objectives of the study which are to assess the methodology's potential for improving collaborative understanding, together with feasibility considerations, i.e., factors that may hinder the adoption or effective use of the methodology in an operational context. Collaborative Understanding was subdivided into three components (Integration of perspectives, Common Understanding, and Collaboration Support) to mirror the design team's targeted areas of support. Note that performance (i.e., end-product quality) is not included here since the objective is to focus

on the support provided by the 12om methodology rather than to assess the team's capability.

Table 2 summarizes the sources of the metrics and their calculation.

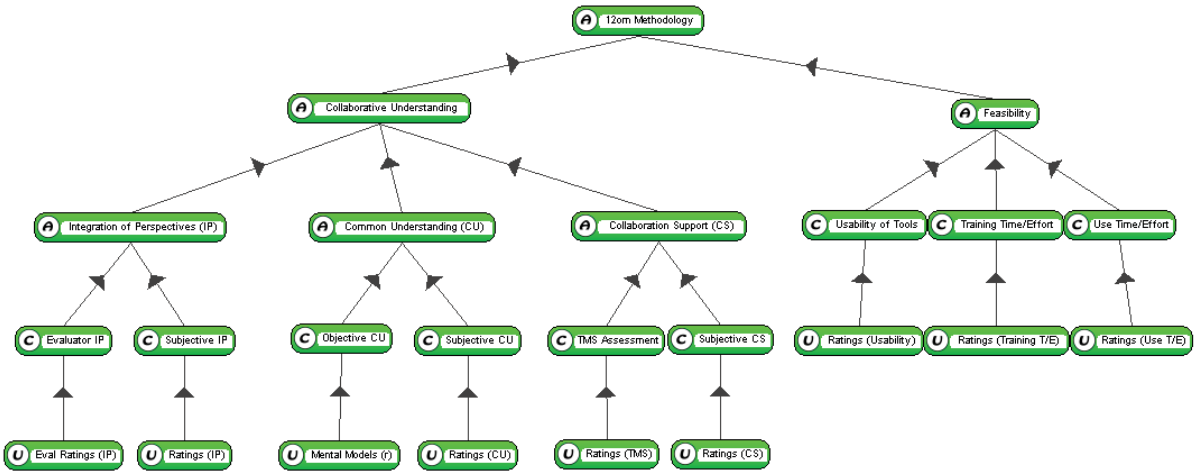


Figure 6: Methodology Assessment Using a MYRIAD Preference Model. U-nodes = metrics, C-nodes = criteria, and A-nodes = Choquet aggregation.

Table 2. Source of metrics and calculation for the overall 12om methodology multi-criteria assessment

Metrics (U)	Questionnaire	Item	Calculation
Evaluator Ratings (IP)	Mission analysis brief evaluation (Appendix J)	Rate the degree of integration (1=Poor, 5=Thorough)	Average score of evaluators for question B2
Ratings (IP)	Methodology assessment questionnaire (Appendix A)	Usefulness for supporting IP (0-10)	Average score for the item
Mental Models (r)	Team factor assessment questionnaire (Appendix D)	Rate the possible impact of each factor on mission achie-vement (-3, 0,+3)	Average correlation of each individual to group (-1 to 1)
Ratings (CU)	Methodology assessment questionnaire (Appendix A)	Usefulness for supporting CU development (0-10)	Average score for the item
Ratings (TMS)	Transactive memory systems (TMS) (Appendix F)	See 10 items in Appendix F (1-5)	Average ratings after reversal of scores for negative questions
Ratings (CS)	Methodology assessment questionnaire (Appendix A)	Usefulness for supporting the collaboration process (0-10)	Average score for the item
Ratings (Usability)	Component assessment questionnaire (Appendix B)	What is the usability (relative to other tools that may be used in this context) of "X"? (0-10)	Average usability ratings of 12om components
Ratings (Training T/E)	Component assessment questionnaire (Appendix B)	How much training [time/effort] (relative to other tools that may be used in this context) is required for	Average training time and effort ratings of 12om components

Metrics (U)	Questionnaire	Item	Calculation
		training on "X"? (0-10)	
Ratings (Use T/E)	Component assessment questionnaire (Appendix B)	How much use [time/effort] (0-10) (relative to other tools that may be used in this context) is required for training on "X"?	Average use time and effort ratings of 12om components

The model includes metrics (U-nodes) at the bottom of the hierarchy which correspond to experimental measures. Criteria nodes (C-nodes) essentially serve to map metrics to a utility function on a 0-1 scale which can be simply linear or capture highly non-linear preference relations. The utility functions mapping the metric values to the degree of satisfaction of each criterion were set to approximate a sigmoid function (see Figure 7). This type of function is generally deemed to provide a more accurate representation of psychological processes than a linear function (Bogacz, Brown, Moehlis, Holmes, & Cohen; 2006; Kilian & Siegelmann, 1996; Kuznar, 2002; Pew, 1969; Sharp, Viswanathan, Lanyon, & Barton, 2012; Wickens & Hollands, 2000).

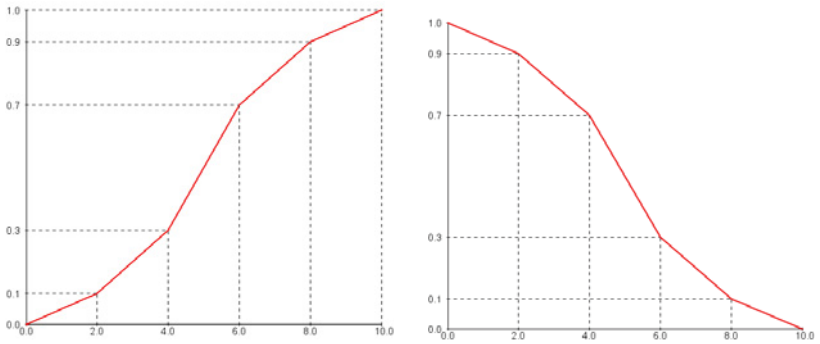


Figure 7: Sigmoid-like utility functions used to represent positive and negative satisfaction curves (x-axis = metric value, y-axis = HFE satisfaction)

Aggregation nodes (A-nodes) allow combining criteria using the Choquet integral method. Conceptually, this means creating an aggregation function that defines the relation between elements on a continuum ranging from purely disjunctive (independent and additive) to purely conjunctive (co-dependent and complementary). The parameters of the Choquet integral are derived automatically based on the constraints (preferences) specified by the modeller. For the present purposes, aggregation elements are assumed to be equally important and the question here was essentially to determine if these elements were to be qualified as independent (additive preferences), complementary, substitutable, or hybrid (partly independent, partly complementary or substitutable).

In the general assessment preference model, an asymmetric hybrid aggregation model, shown in Figure 8, was used to define the top node "12om Methodology", in order to properly capture that collaborative understanding is a key objective in itself, but that feasibility is important only when in conjunction with a good degree of collaborative understanding). For example, a high feasibility but with a low collaborative understanding will yield a low result (i.e. the red portion which has 50% of the weight will be low, and the

purple portion which has the other 50% of the weight will be low as well because the complementarity means it is the lowest of the two values that will be considered for this portion), In practice, this means that if collaborative understanding is 40% and feasibility is 80%, the overall satisfaction will be 40%. Conversely, if collaborative understanding is 80% and feasibility is 40%, overall satisfaction will be (80 X 0.5) + (40 X 0.5), hence 60%.

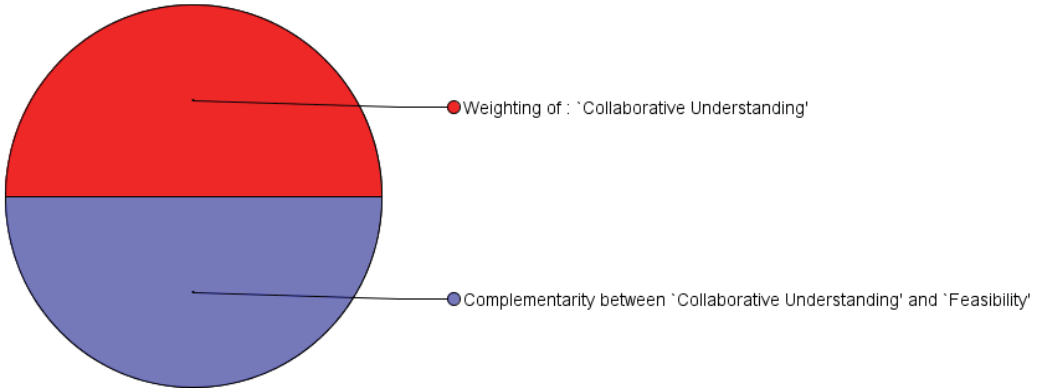


Figure 8: Hybrid aggregation (used to define the top node in the Component Assessment Model)

Collaborative Understanding was defined using a purely complementary relation of its three elements to capture their high degree of interdependence, as shown in Figure 9. Indeed, the goal of the 12om methodology is to jointly support these three dimensions, and a balanced support across these variables will be preferred to a less balanced combination of greater and lower support values.

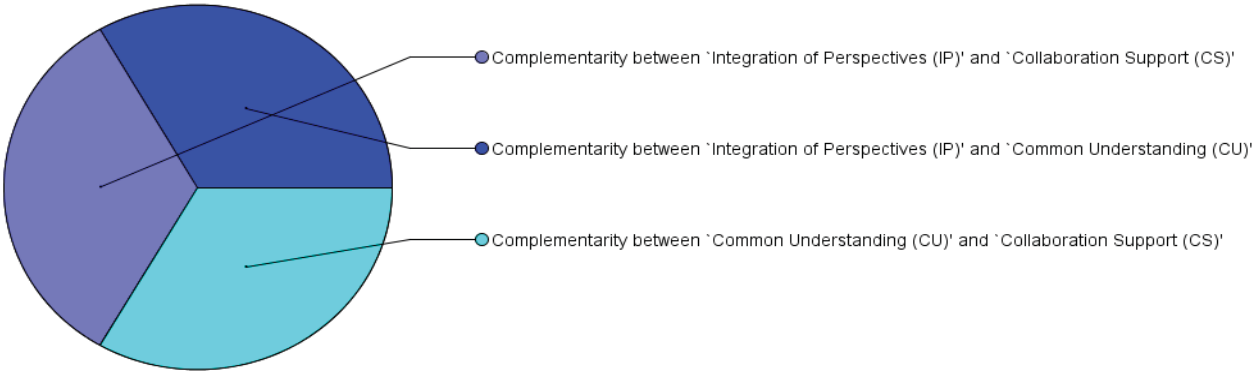


Figure 9: Conjunctive aggregation (defining "Collaborative Understanding" in the General Assessment Model)

A symmetric hybrid model was used to define the Feasibility aggregation, shown in Figure 10. In this case each element is partially independent and partially complementary, in order to represent the synergy between these elements in supporting Feasibility.

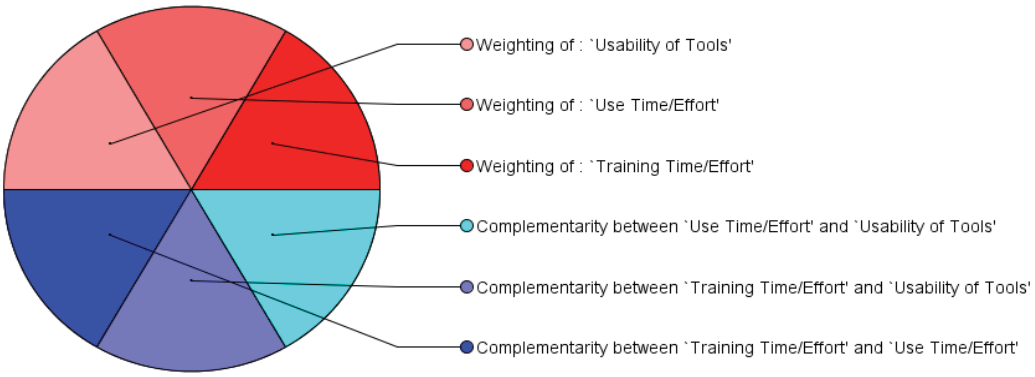


Figure 10: Hybrid aggregation (defining "Feasibility" in the General and Component Assessment Models)

For example, if two criteria are at 100% and one at 50% the overall satisfaction will correspond to (100 X 16.66) + (100 X 16.66) + (50 X 16.66) + (50 X 16.66) + (50 X 16.66) + (50 X 16.66), hence 66.66%.

Finally, the level 3 aggregations in the general assessment model (Integration of perspectives, Common Understanding, and Collaboration Support) are defined using a simple additive relation between their lower-level respective elements.

Individual Component Assessment

The individual component assessments included in the MYRIAD analysis focused on the twelve following elements: Common glossary, OPP Handbook, Team building, Conceptual Diagrams (Individual), Conceptual Diagrams (Collaborative), Creation of the common vocabulary, Creating views using filters, Querying conceptual diagrams using filters, Sharing conceptual diagrams, Cross-Impact Method, Op design tool, and WoG MA Brief Template. Figure 11 shows the preference model used for the individual component assessments. The top of the hierarchy remains similar to the general assessment, with the right side (feasibility) untouched, but with a broader assessment on the left side to get a better sense of the support provided within the present task context, hence the two new nodes on the left, namely *Impact on Taskwork* and *OPP Support*. *Impact on Taskwork* is based on a metric derived from the task-to-tool mapping, more specifically the weight given by the team as a whole to rate the relative impact of each component for performing the task (which can range from 0 – outright exclusion of the component, followed by 1 – inclusion, and up to 4 which corresponds to three stars). Lower impact values were qualified as principally serving as reference material, while higher impact ratings were qualified as aiding in the accomplishment of the cognitive work. The *OPP Support* criterion was based on the average rating in the component assessment questionnaire about the component's degree of support for activities and products of OPP. Table 3 summarizes the sources of the metrics and their calculation.

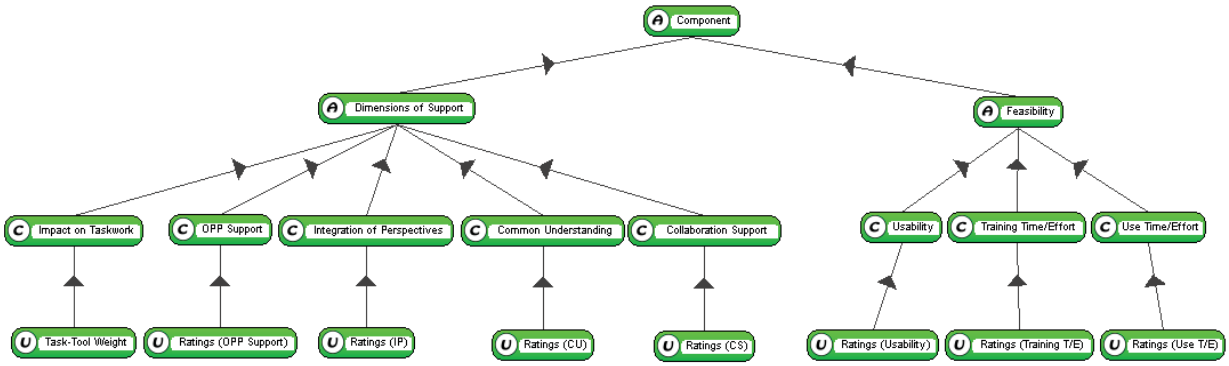


Figure 11: Methodology Component Analysis Using a MYRIAD Preference Model.

Table 3. Source of metrics and calculation for the components multicriteria assessment

Metrics (U)	Questionnaire	Item	Calculation
Task-tool weight	Task-to-tool mapping (Appendix H)	Final synthesis of task-to-tool mapping (see Figure 40). From 0 (exclusion), to 4 (three stars).	Collective team rating of impacts of each component in the task-to-tool mapping exercise (0-4)
Ratings (OPP Support)	Component assessment questionnaire (Appendix B)	Rate the usefulness for supporting OOP (0-10)	Average score for the item "supporting activities and products of OPP"
Ratings (IP)	Component assessment questionnaire (Appendix B)	Rate the usefulness for supporting IP (0-10)	Average score for the item "supporting integration of perspectives"
Ratings (CU)	Component assessment questionnaire (Appendix B)	Rate the usefulness for supporting CU development (0-10)	Average score for the item "supporting common understanding"
Ratings (CS)	Component assessment questionnaire (Appendix B)	Rate the usefulness for supporting the collaboration process (0-10)	Average score for the item "supporting collaboration"
Ratings (Usability)	Component assessment questionnaire (Appendix B)	What is the usability (relative to other tools that may be used in this context) of "X"? (0-10)	Average usability ratings for the component
Ratings (Training T/E)	Component assessment questionnaire (Appendix B)	How much training [time/effort] (relative to other tools that may be used in this context) is required for training on "X"? (0-10)	Average training time and effort ratings for the component.
Ratings (Use T/E)	Component assessment questionnaire (Appendix B)	How much use [time/effort] (0-10) (relative to other tools that may be used in this context) is required for training on "X"?	Average use time and effort ratings for the component.

A key difference in the individual component assessment model is that the "Dimensions of Support" aggregation node is defined using a purely additive relation, since individual components are not required to simultaneously address all dimensions of support (i.e., components can be useful even if the breadth of their impact is limited). Figure 12 shows the particular relation used to define the "Dimensions of Support" aggregate variable.

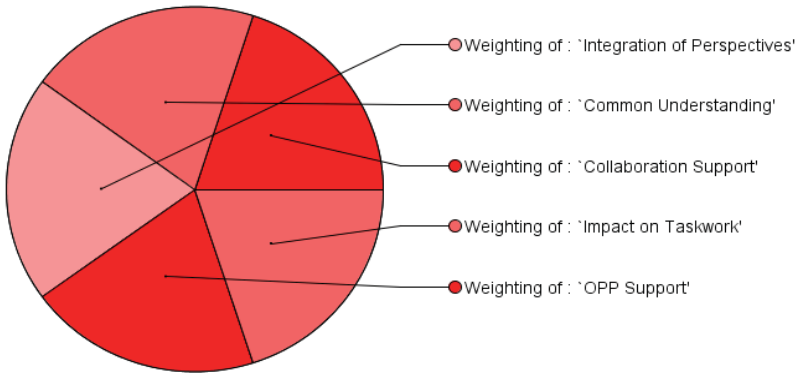


Figure 12: Additive aggregation (defining "Dimensions of Support" in the Component Assessment Model)

The MYRAD preference model described here constitutes an initial assessment mainly based on the analyst team's perspective. This model may be improved or fine-tuned in the future through interactions with SMEs and stakeholders.

2.3.2 Performance Analysis

In addition to "support to collaboration", "integration of different perspectives", and "ability to develop a common understanding", 12om aims to increase performance of teams as a whole, notwithstanding other dimensions. Assessing performance represents an important challenge in the context of the second LOE since it cannot be directly compared to the first LOE given the large amount of differences between the two experiments. For instance, LOE#2 assessed the performance of a 5-member planning team whereas LOE#1 evaluated 4-member planning teams. The scenario was also different between the two experiments. Furthermore, the time allocated to complete the mission analysis brief was far from equivalent across experiments. Finally, the second LOE covered a broader spectrum of the OPP tasks, including the initial steps carried out for COA development. However, several means were deployed to capture a valid portrait of performance in this context, if only for descriptive purposes and as a proof of concept test that performance can be good when using the 12om methodology. Measures of performance (MoP) included NASA-TLX and mission awareness rating scale, and measures of effectiveness (MoE) consisted of the SME observers' evaluation form.

2.3.2.1 Measures of Performance

NASA-TLX

Workload is a critical element associated with human performance. Although workload is often treated as task-dependent, some authors argue that it is task-independent; mostly determined by factors such as cognitive abilities and tools (e.g., Gonzalez, 2005). New tools may potentially reduce the participants' level of workload to an optimal level. To investigate that hypothesis, the NASA-TLX (Hart & Staveland, 1988) measure of workload was used. NASA-TLX also provides a level of workload scores for each participant.

Five dimensions of workload were assessed during the experiment; the mean score of each dimension are reported:

1. *Mental demand*. How much mental and perceptual activity was required? Was the task easy or demanding, simple or complex, exacting or forgiving?
2. *Temporal demand*. How much time pressure did you feel due to the rate or pace at which the task or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
3. *Performance*. How successful do you think you were in accomplishing the goals of the task set by the experimenter? How satisfied were you with your performance in accomplishing these goals?
4. *Effort*. How hard did you have to work mentally to accomplish your level of performance?

5. *Frustration*. How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

The NASA-TLX questionnaire is shown in Appendix E.

Mission Awareness Rating Scale (MARS)

Mission awareness was assessed with the Mission Awareness Rating Scale (MARS; Matthews & Beal, 2002). The MARS is a self-rating assessment technique designed specifically for use in the assessment of situation awareness (SA) during a military exercise, and provides information on the participants' capacity to acquire and maintain SA. The first level of SA is to perceive the status, attributes, and dynamics of relevant cues in the environment. The second level is to comprehend the situation based on a synthesis of Level l cues. Comprehension goes beyond simply being aware of the cues that are present to include an understanding of the significance of those cues in the context of the actual task. Finally, the third level of SA builds upon level 1 and level 2 SA to project future actions and their effects in the environment. MARS comprises two separate sets of questions: one assessing the ability to acquire SA and another to assess the difficulty of maintaining SA. The MARS is available in Appendix G.

Process evaluation

The process evaluation questionnaire was handed out to SME evaluators in order to assess the performance of the integrated planning team with regard to their process, independently from the quality of their mission analysis brief. The questionnaire was composed of 31 items, but was reduced to 28 after removing items related to requests for information as none was produced during the scenario. SMEs had to rate each item on a scale ranging from 1 (not at all) to 5 (very much) with 3 being average. The average of SME ratings per item is reported. The process evaluation questionnaire is shown in Appendix I.

2.3.2.2 Measures of Effectiveness

An evaluation form was handed out to the SME observers so they could rate the performance of the WoG planning team with regard to the quality of their MA brief. Items concerned the level of detail of the brief, the quality of the brief and if the brief met expectations of the TFC and the ROC.

Items used to extract MoE were in Appendix J, items A1 to A6. Each item were averaged across evaluators and reported individually.

2.3.3 Common Understanding

A key aspect of the 12om methodology is to foster the development of a common understanding, or team mental model, of a complex situation. Team mental models enable a group of individuals to communicate effectively and accomplish complex tasks by acting in a coordinated manner (Hsu, Chang, Klein, & Jiang, 2011). Research has shown that the quality of common understanding is related to the likelihood of having an effective team (Cannon-Bowers & Salas, 2001; Mohammed & Dumville, 2001). In this study, common understanding was inferred through a measure of mental model cohesion, which refers to

the degree at which teammates agreed about the importance of factors associated with the mission.

Common understanding (cohesion) was measured through the *individual factors identification* (Appendix C) and *team factor assessment* (Appendix D) questionnaires. These questionnaires were administrated in two phases. The first phase required participants to individually identify factors that they believed were associated with mission achievement. This set of factors is referred to as "intra-mission achievement factors". Participants also had to identify the top three factors associated with mission achievement from the perspective of the other team members. For instance, J5 had to identify the factors he believed J25, J5Ops, J5Dev and J5Gov though were important for mission achievement. This set of factors is referred to as "inter-mission achievement factors". Intra-mission achievement factors were then merged into a single list by the experimenters to avoid overlap. In the second phase, participants, as well as the ROC and the TFC, were given the merged list of factors and were required to rate the impact of each factor on mission achievement following instructions in Figure 15:

Rate the possible impact of each factor on a 7-point Likert scale.

0 = ***No*** impact of the factor on mission achievement

+3 = Strong ***positive*** impact on mission achievement

-3 = Strong ***negative*** impact on mission achievement

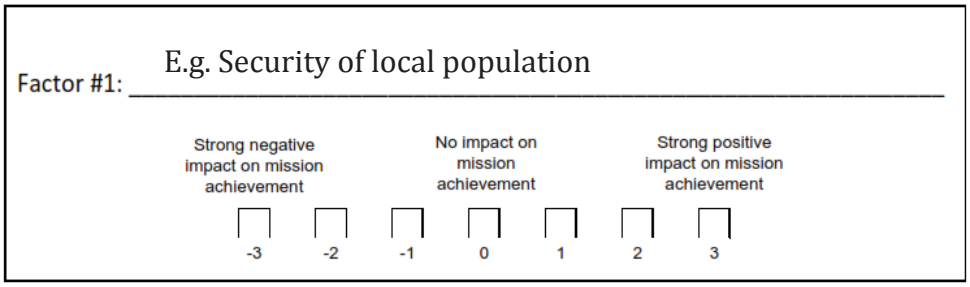


Figure 13. Example of the "team factor assessment questionnaire".

Cohesion was estimated by correlating ratings of the team members together: high correlation coefficients being associated with high cohesion and vice versa.

Inter-mission achievement factors were also analysed. Each member of the team had to identify three factors important for each of the other members. The three factors were compared against the 10 intra-mission achievement factors. Whenever one of the inter-mission factors corresponded to an intra-mission factor, it was calculated as a "hit". The average of "hits" is reported. Theoretically, the maximum average of hits is 3, which would suggest a very good understanding of the other team members' mental model. Conversely, an average of 0 would suggest a lack of knowledge about the mental model of the other

members of the planning team. Finally, the number of overlapping intra-mission factors was also calculated. This measure is another indicator of the common understanding of the team.

2.3.4 Functional Gap Analysis

A functional-gap analysis was performed in order to identify critical tasks/processes in the orientation phase of the CF OPP in relation to the key support capabilities of the 12om methodology. Functional gap analysis involves the preparation of a preliminary representation of the CF OPP tasks and sub-tasks, and 12om components by the research team. Figure 14 provides a hypothetical example of a mapping between CF OPP orientation tasks and the capability and functionality of a decision support tool. The goal of the analysis is to physically map each system capability and functionality with the CF OPP task(s) that it is designed to support. This task-to-tool mapping allows the analyst to assess what critical orientation tasks are supported by the tools, and what tasks are not supported (e.g. see arrow pointing a specific node in Figure 14).

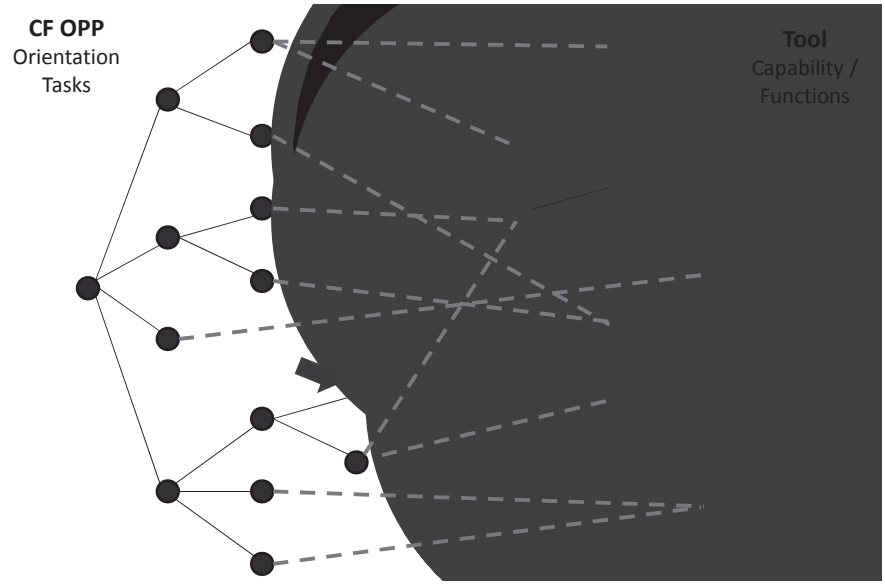


Figure 14. Hypothetical example of a mapping between CF OPP orientation tasks and the capability and functionality of a decision support tool.

A large poster representation was presented to participants (see Appendix H). On the left side of the poster was presented a hierarchical decomposition of the main tasks associated with the OPP initiation, MA, and initial COA development phases. On the right side, the 12om components were represented. Participants were asked to individually perform the functionality mapping on a printed sheet. For each component, they had to draw links to the tasks that it supported. Afterward, participants were asked to collectively perform this mapping on the main large-scale cardboard, resolving any differences along the way. Along the mapping process, they were invited to add or remove any task or component that they felt was missing or useless. Finally, participants were asked to rate the impact of each components in terms of support, on a scale ranging from 0 to 4. Overall, the analysis

informs on the level of support of each component to OPP both in terms of flexibility and impact. In addition, the analysis allows the identification of unsupported or insufficiently supported sub-tasks of the OPP.

2.3.5 Team Dynamics

Team dynamics is the umbrella term that groups together the analyses aiming to describe the interactions between the WoG team members. It can be useful for assessing the collaboration between team members, the integration of different perspectives, the development of a common understanding, and identify ways to improve individual components of the 12om methodology. Team dynamics analyses include several sub-analyses; however, the input data for all analysis was collected through video/audio recordings made with the webcams installed in the room. This data allowed us to carry out social network analysis and communication content analysis.

2.3.5.1 Social Network Analysis

Social network theory (e.g., Wasserman & Faust, 1999; Watts, 1999; Watts & Strogatz, 1998) is used as a tool to investigate the organizational structure of the team of planners for which the connections between entities can reveal a great deal about the pattern of information exchange (who talks to who).

The first step is to collect data that describe relationships between entities in terms of communications, relationships, or transactions to create a matrix. Frequency of verbal communications was extracted from the recordings made during the experiment. The data matrices were then used to construct a social network graph; a visualization illustrating the connections between entities from which one can readily observe structures and relationships (see example on Figure 15). Furthermore, mathematical analyses were applied to the matrix itself to quantify aspects of the network numerically.

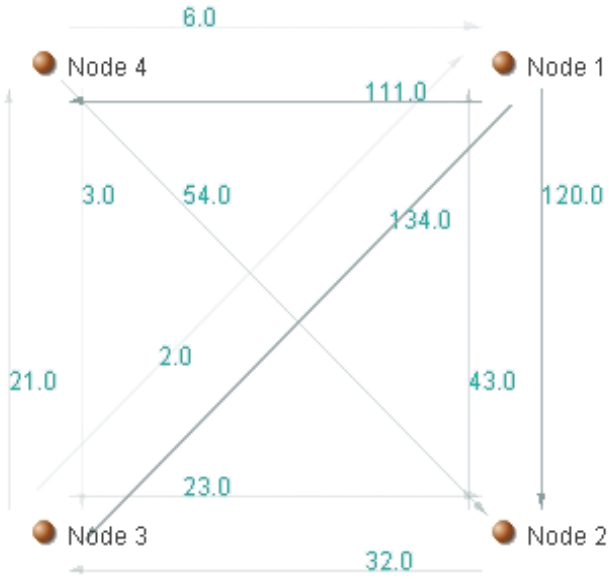


Figure 15: Example of a social network graphical representation. Each node represents an individual. Arrows between nodes represent the existence of a link between individuals. The direction of the arrow shows the flow of information within the network. The number beside the arrow indicates the frequency of interaction associated with the link. The thickness and opaqueness of the arrows indicate the importance of the link in relation to the other links within the network. For instance, we can observe that "Node 1" transfers a lot of information to "Node 3", but that the opposite is not true.

Three analyses were conducted on the communication interaction data (i.e., frequency of communications between nodes) that was collected during the experiment – emission degree, reception degree, and sociometric status (e.g., Benta, 2005). A communication refers to the transfer of a single idea from an agent to another. A sentence can be composed of many ideas and will consequently be coded as multiple communications. Emission degree is a metric describing communications at the individual level. The emission degree of a node is the sum of all values corresponding to the edges originating in that node. The reception degree of a node is the sum of all values corresponding to the edges incident to that node. Sociometric status is a measure describing communications at the individual level. The sociometric status of a node is the sum of its reception and emission degrees, relative to the number of all other nodes in the network.

2.3.5.2 Content Analysis

In addition to the social network analyses, a subset of the communications was further analysed to extract its content. The subset of communications was determined in collaboration with the DRDC scientific authority. The following time periods were retained as they were undisturbed by training and/or use of specific components:

Day 1 pm – One hour of undisturbed mission analysis, during scenario read in.
Day 2 am – One hour 45 minutes of undisturbed mission analysis, during scenario read in.
Day 3 am – Twenty-one minutes of undisturbed mission analysis, after collaborative knowledge representation training.

Day 3 pm – One hour 5 minutes of undisturbed mission analysis, before the OP Design tool presentation.
Day 4 am – Forty-five minutes of undisturbed mission analysis, before mission analysis brief.

The choice of selecting the undisturbed phases was motivated by the fact that the "disturbed" phases were artificial in their nature because they included training/commenting on tools and other activities that are not normally directly associated with OPP. Within these time windows, each communication was categorized as one of the following content types: (a) security, (b) development, (c) governance, (d) process, (e) team building, and (f) others. Table 4 summarizes the main keywords that served as guidelines to categorize each of the communications.

Table 4. Keywords used to guide communications categorization.

Security	Development	Governance	Process	Team building	Others	Inaudible
Military; Safety; Protection; Security; Terrorism; Al-Shabaab; Tribal conflicts; Support; Lines of communication; Stability/instability; Clan rivalry	Assist; Facilitation; support; resources; food; camp; health; ministry of health; health cluster; backing; NGO; IDP; vaccination; immunization; health system; public health; quarantine.	Social control; law making; legislation; destabilization; protests; government; authority; regime; power; mediation; legitimacy; FGS	Coordination; planning; synchronization; allocation; deliverable; product; brief; mission analysis; mission statement; course of actions	Personal; socialization; breaks; doughnuts	Incomplete sentences; general acknowledgment not related to specific communication; others	Inaudible

The content analysis is used to show the relative importance of each of these categories during the planning cycle. Moreover, the rate of communication by line of operations (i.e., security, governance, and development) was calculated for each of the aforementioned time periods, allowing for a comparison of their relative importance through time.

2.3.5.3 Transactive Memory System

The TMS questionnaire, specifically aiming to assess team dynamics, was administered to the participants (see Appendix F). TMS is the cooperative division of labour for learning, remembering, and communicating relevant team knowledge. According to TMS theory, group members divide the cognitive labour for their tasks, with members specializing in different domains. Members rely on one another to be responsible for specific expertise such that collectively they possess all of the information needed for their tasks. Although a

TMS is a group-level phenomenon, it exists as a function of the structure, content, and credibility of members' individual knowledge, implying that it may be appropriately measured at the individual level. The conceptual descriptions of a TMS suggest it is multidimensional, manifesting in specialized knowledge, credibility perceptions, and coordination processes.

TMS subscales used in the current study differentiate members' beliefs about the reliability of other members' knowledge (i.e., *credibility*), and effective, coordinated knowledge processing (i.e., *coordination*). To measure TMS (and its subscales), a self-report questionnaire developed by Lewis (2003) was used. Average ratings in terms of credibility and coordination are reported.

2.3.6 Qualitative assessment

A fair amount of qualitative data was collected during the experiment. Sources for qualitative data included video and audio recordings of the experiment, including the task-to-tool mapping exercise and the focus group discussions. Other sources were the comment sections of the component assessment questionnaires, mission analysis process evaluation, and mission analysis brief evaluation (respectively available in Appendix B, Appendix I, and Appendix J).

3 Results

This section reports the main results of the second LOE. We first report the results of the 12om methodology assessment. We present an assessment from different perspectives, both on the methodology as a whole and by component. Depending on the analysis, the focus of the assessment is either on the costs and benefits associated with the use of the methodology, or on the level of achievement of its main objectives, namely to facilitate the integration of perspectives, ease the collaboration process, and increase common understanding. We then present results pertaining specifically to the performance of the team with regard to mission objective (which was to produce a MA brief). Finally, we present a set of results aiming to provide additional insight about the impacts of the 12om methodology including outcomes in terms of flexibility, impact on task work, and effects on team dynamics.

3.1 12om Methodology Assessment by Criterion

The methodology assessment concerns all analyses directly related to the evaluation of the 12om processes/tools. When relevant, results are reported separately for SME observers and WoG planning team. Interpretation of the results in this section prioritizes the ratings of the WoG planning team as they directly interacted with the 12om methodology components as opposed to indirect observations from the SME observers. We assume the direct manipulation/use of the components helped raters appreciate more adequately the strengths and weaknesses of the methodology components. Discrepancies observed between ratings of the SME observers and of the WoG team may originate in part from the reduced validity of ratings of the SME observers and may not reflect "real" differences. For this reason, discrepancies between groups of participants should be interpreted with caution.

3.1.1 Overall 12om Methodology Assessment

A questionnaire aiming to assess the level of support of the 12om methodology on key dimensions was handed out to the WoG planning team members (n = 5) and SME observers (n = 4) (Appendix A). All items of the questionnaire were considered for the analysis. The key dimensions were:

- Integration of different perspectives;
- Common understanding; and
- Collaboration.

Mean ratings and standard errors are reported in Figure 16 by evaluators (WoG team and observers) and key dimensions.

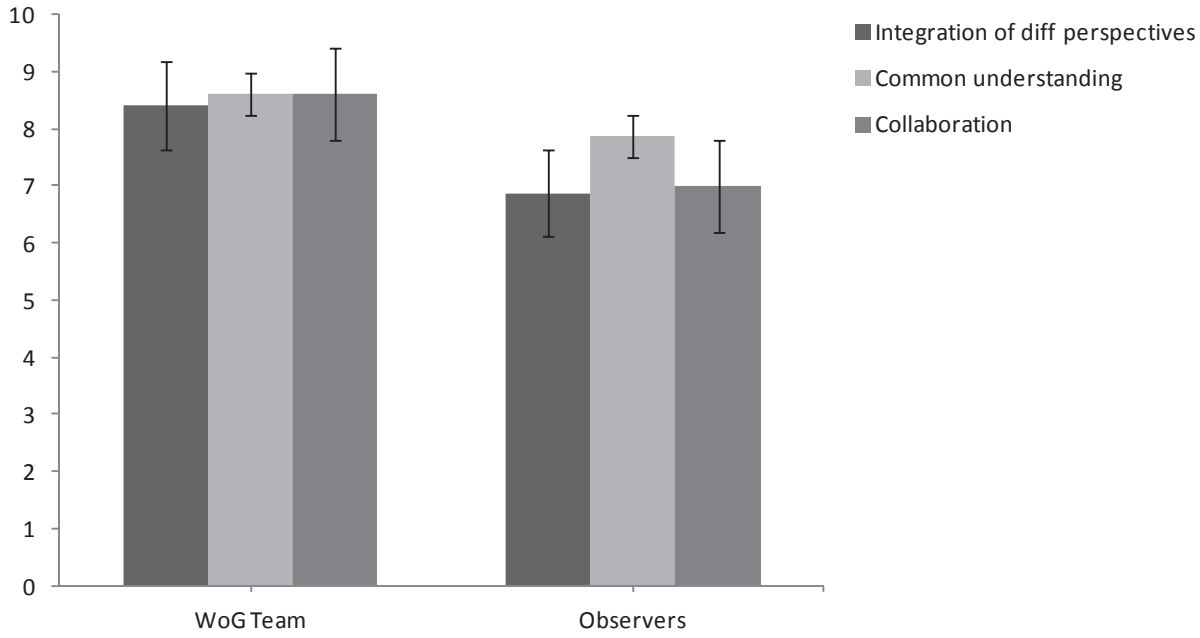


Figure 16. Average ratings of the perceived level of support provided by the 12om methodology by key dimensions of support and evaluators. Error bars represent standard error.

Overall, the 12om methodology was rated very favourably in terms of support to the key dimensions. Moreover, there were no dimensions significantly under rated compared to the others. Participants who interacted with the components directly (WoG team) rated 12om support of key dimensions higher compared to the observers' team.

3.1.2 Assessment of 12om Components

Each of the components (or sub-components) was assessed through a series of questions submitted to the members of the WoG team and SME observers. Results are reported by component for each of these questions in Figure 17 for usefulness and in Figure 18 for training time and effort for the planning team ratings. SME ratings were more variable and contained more missing values. The reason for this is that they did not interact directly with the components and may sometime have had a limited visibility on the planning team. For this reason, the ratings of the SME observers to these items are reported in Appendix K.

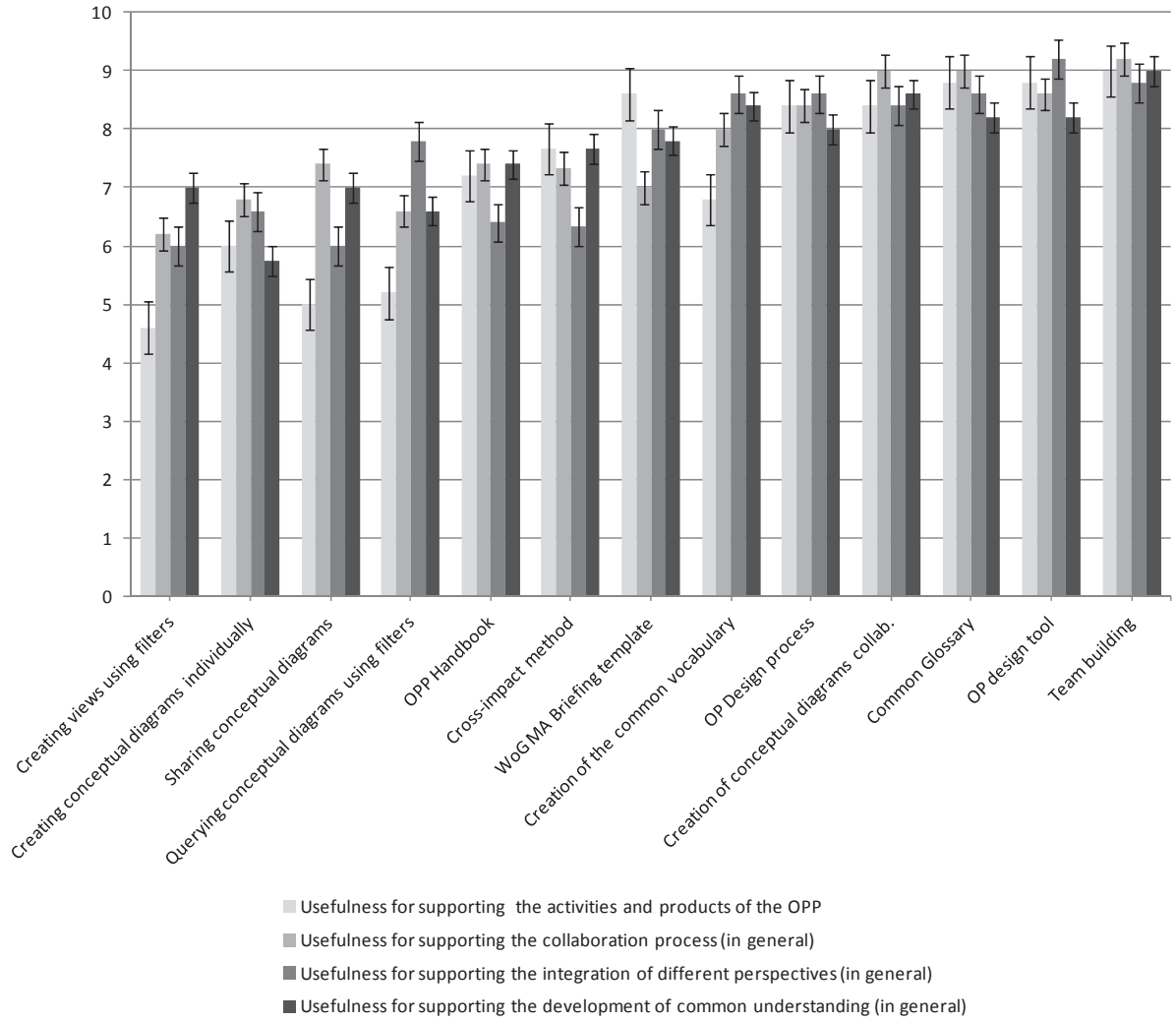


Figure 17: Average ratings of the planning team by usefulness item and component.

Results show no major differences on the four items related to usefulness across all components. However, the more specialized components associated with the IMAGE v3 tool appear to be considered a little less useful in supporting OPP than in supporting collaboration, integration of different perspectives and common understanding. This could be interpreted as being a sign that these specialized components are generic in their application, which is reasonable. Indeed, the creation of views using filters or sharing conceptual diagrams applies to many different contexts, not necessarily OPP.

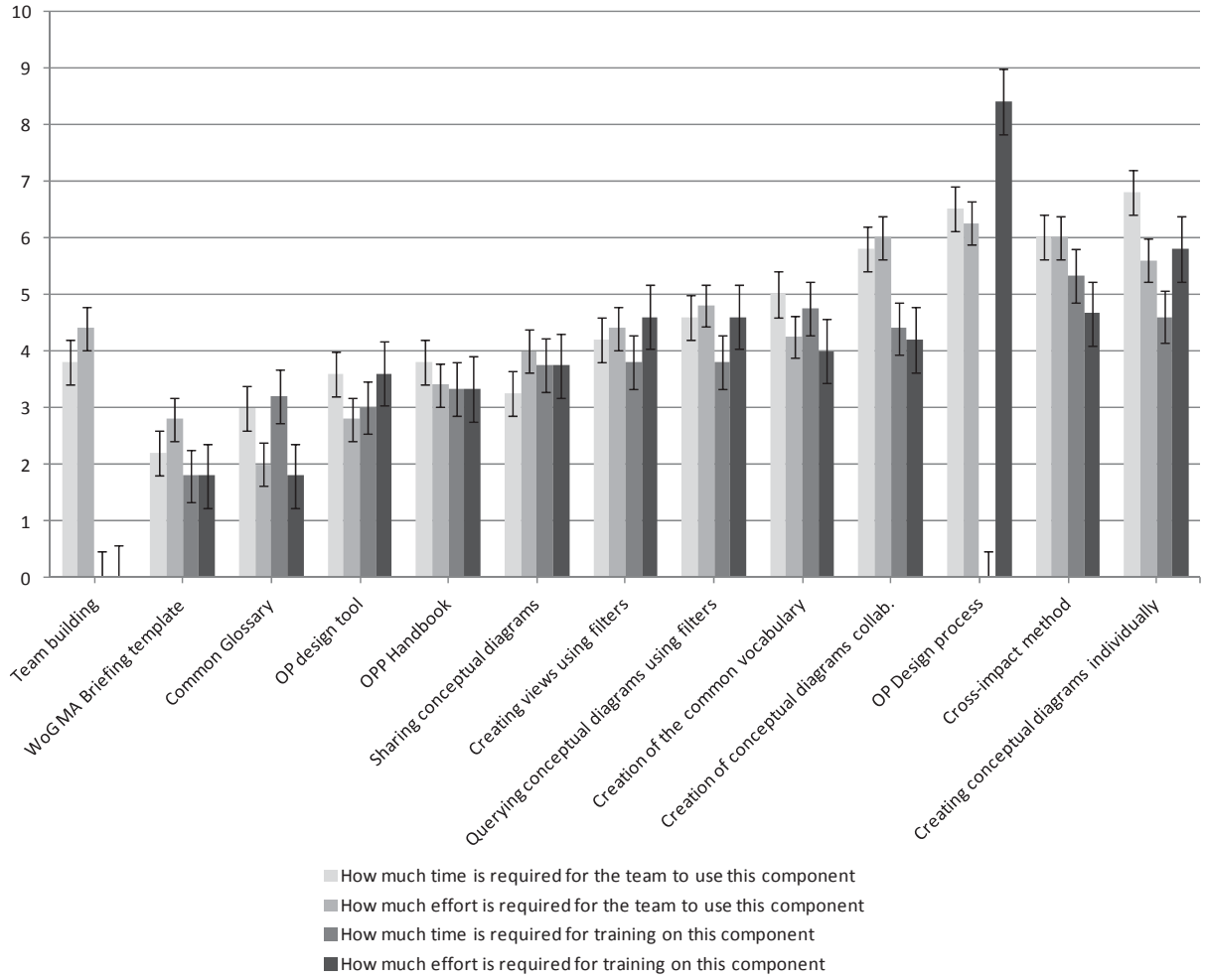


Figure 18: Average ratings of the planning team by time and effort item and component.

Results show that in some cases, training time/effort and use time/effort are comparable within a given component. They also show that apart from team building and handover procedure, all components would require at least some training.

A series of mean ratings representing each of the following concepts was calculated: (a) usefulness of the component, (b) time/effort required to use and train the component, (c) usability of the component, and (d) cost/benefit function. The mean rating was calculated separately for the WoG team members and the SME observers.

Usefulness

Figure 19 represents average ratings of observers and WoG team with regard to the usefulness of each component and sub-component assessed during LOE #2. Components are arranged from left to right in ascending order of ratings (by WoG team). The ratings can theoretically vary between 0 (very low usefulness) and 10 (very high usefulness).

Overall, all components were rated to be useful. Only two components were rated below 6/10 by the SME observers (i.e., OP design tool and creation of the common vocabulary),

and none was rated below 6/10 by the WoG team. Important discrepancies were observed between the ratings of the SME observers and the ones of the WoG team. From the WoG team's point of view, the less useful components were mostly specific sub-components of IMAGEv3. Being very specific, these sub-components may have failed to demonstrate their potential adequately given the constraints associated with the LOE (notably time constraints). Interestingly, collaborative creation of conceptual diagrams was rated higher in terms of usefulness when compared to individual creation of conceptual diagrams. From the SME observers' point of view, the usefulness of the team building procedure, OP design tool and the creation of common vocabulary were rated much lower than by the WoG team.

Time/Effort

Figure 20 represents average ratings of observers and the WoG team with regard to the time and effort required for using and training each component and sub-component assessed during LOE #2. Components are arranged from left to right in ascending order of ratings (by WoG team). The ratings can theoretically vary between 0 (very little time/effort required) and 10 (a lot of time/effort required).

General time and effort required to use the components was low, even though there is place for improvement. Unsurprisingly, WoG MA briefing template, common glossary, and OPP handbook were among the components requiring the lowest amount of time and effort (according to the WoG team ratings). Conversely, OP design process, creation of conceptual diagrams individually, and cross-impact method were rated as requiring a lot of time and effort. The contrast between the ratings of time and effort required by OP design process versus OP design tool suggests that the tool is not a burden to the application of the process since it doesn't require much additional time to use and/or train. Similarly to the ratings of usefulness, WoG team and SME observers did not fully agree on the time and effort required to use the components. The most important difference concerned the creation of a common vocabulary.

Usability

Figure 21 represents average ratings of observers and the WoG team, by component, on the usability, which was assessed during LOE #2. Components are arranged from left to right in ascending order of ratings (by WoG team). The ratings can theoretically vary between 0 (lower usability than similar tools) and 10 (higher usability than similar tools), 5 being a neutral point (equivalent usability than similar tools).

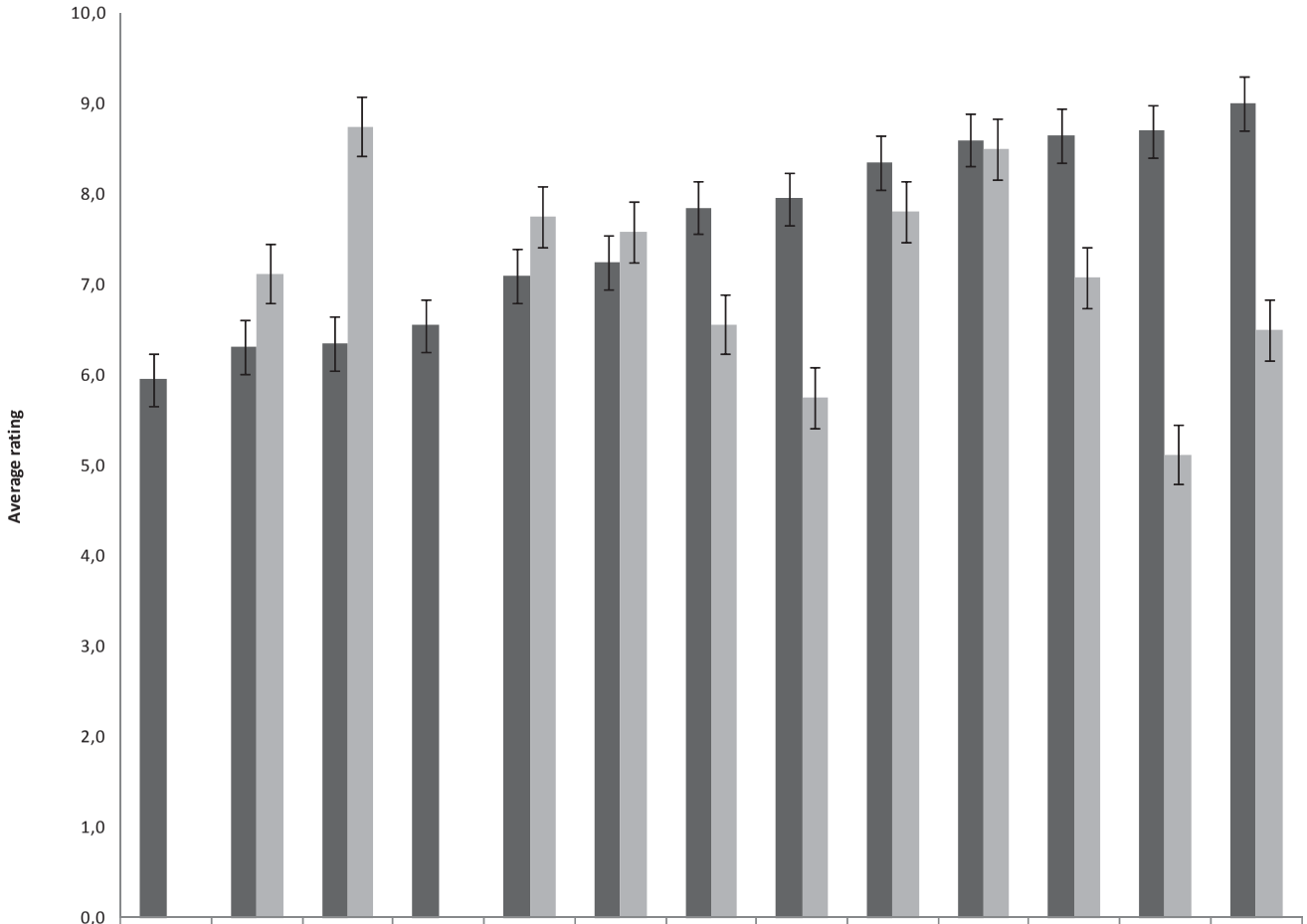
Only a few sub-components were rated as being characterized by lower or equivalent usability than similar tools in this context: i.e. *sharing conceptual diagrams, querying conceptual diagrams using filters, and creating views using filters*. All these sub-components are related to IMAGEv3, however, they do not represent the "essential" features of IMAGEv3. For instance, creation of conceptual diagrams (both individually and collaboratively), were rated very high in terms of usability. Taken together, these results suggest that the core features of the IMAGEv3 tool are valuable for most users, but that the more complex, specialized features should be trained further or supported by a specialized facilitator. An alternative explanation of the lower rating is the lack of time spent on these

components during the case study. “Sharing conceptual diagrams” (which is an easy one-click procedure) was probably also associated with validating graphs (the requirements to define all terms used in the graphs, which is much more labor intensive). All the other components were rated more usable than similar tools used in this context in terms of usability, suggesting that they are mature in that respect. However, there were large discrepancies observed between the two groups of evaluators. OP design tool and team building procedure, for instance, were rated having much lower usability by the SME observers when compared to the ratings afforded by the WoG team.

Costs/Benefits

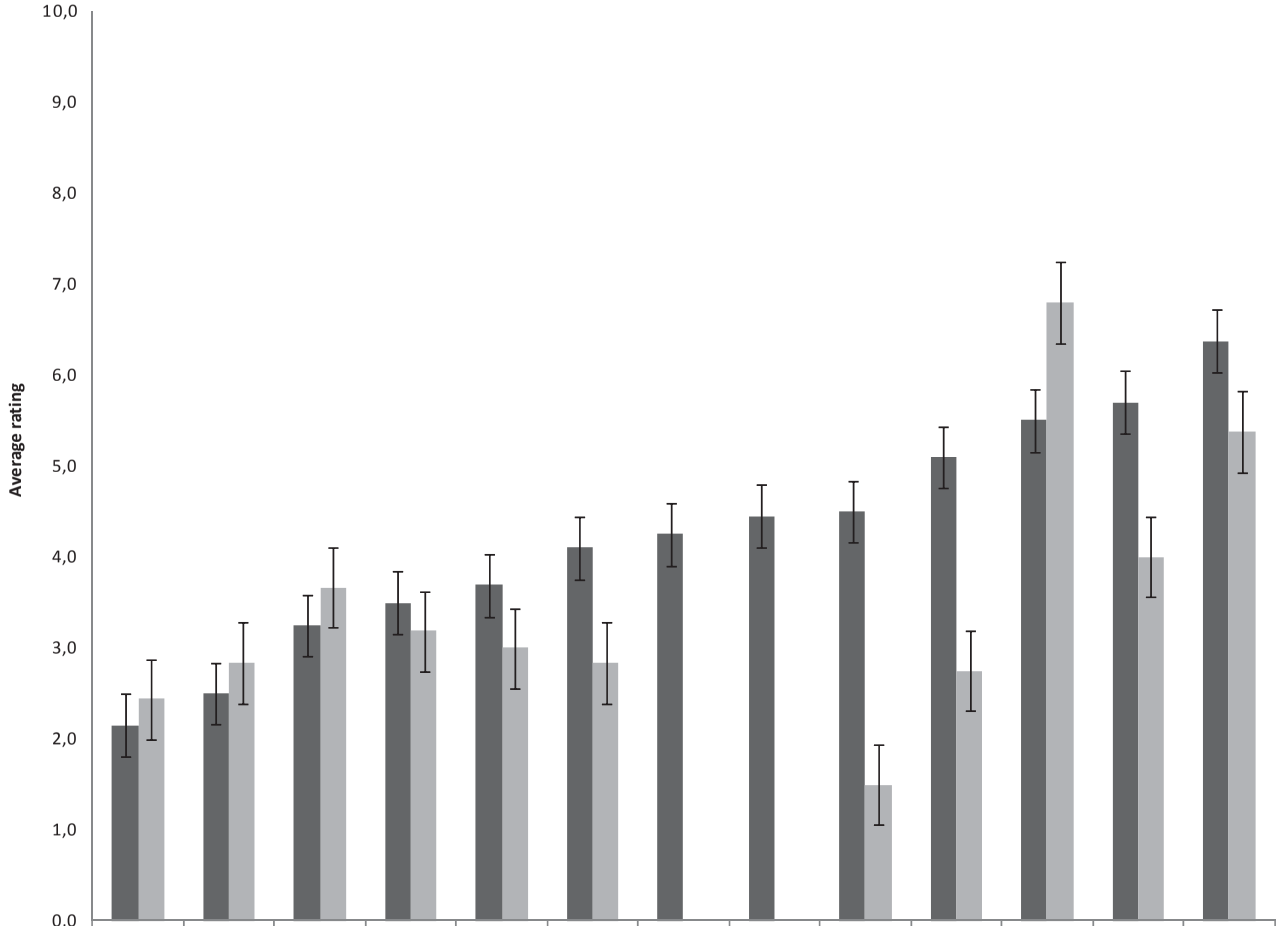
Figure 22 represents average ratings of observers and the WoG team, by component, on the costs/benefits ratio. Components are arranged from left to right in ascending order of ratings (by WoG team). The ratings can theoretically vary between 0 (costs significantly outweigh the benefits) and 10 (benefits significantly outweigh the costs), 5 being a neutral point (costs are equal to benefits).

Most components and sub-components were rated as having benefits that outweighed the costs. Again, the lowest ratings were associated with IMAGEv3 tool sub-components: sharing conceptual diagrams and querying conceptual diagrams using filters. This result suggests that the more complex and specialized features of the IMAGEv3 component may represent a greater challenge of implementation when compared to its core functionalities. For instance, the benefits of the creation of *conceptual diagrams collaboratively* sub-component were rated well above its costs.



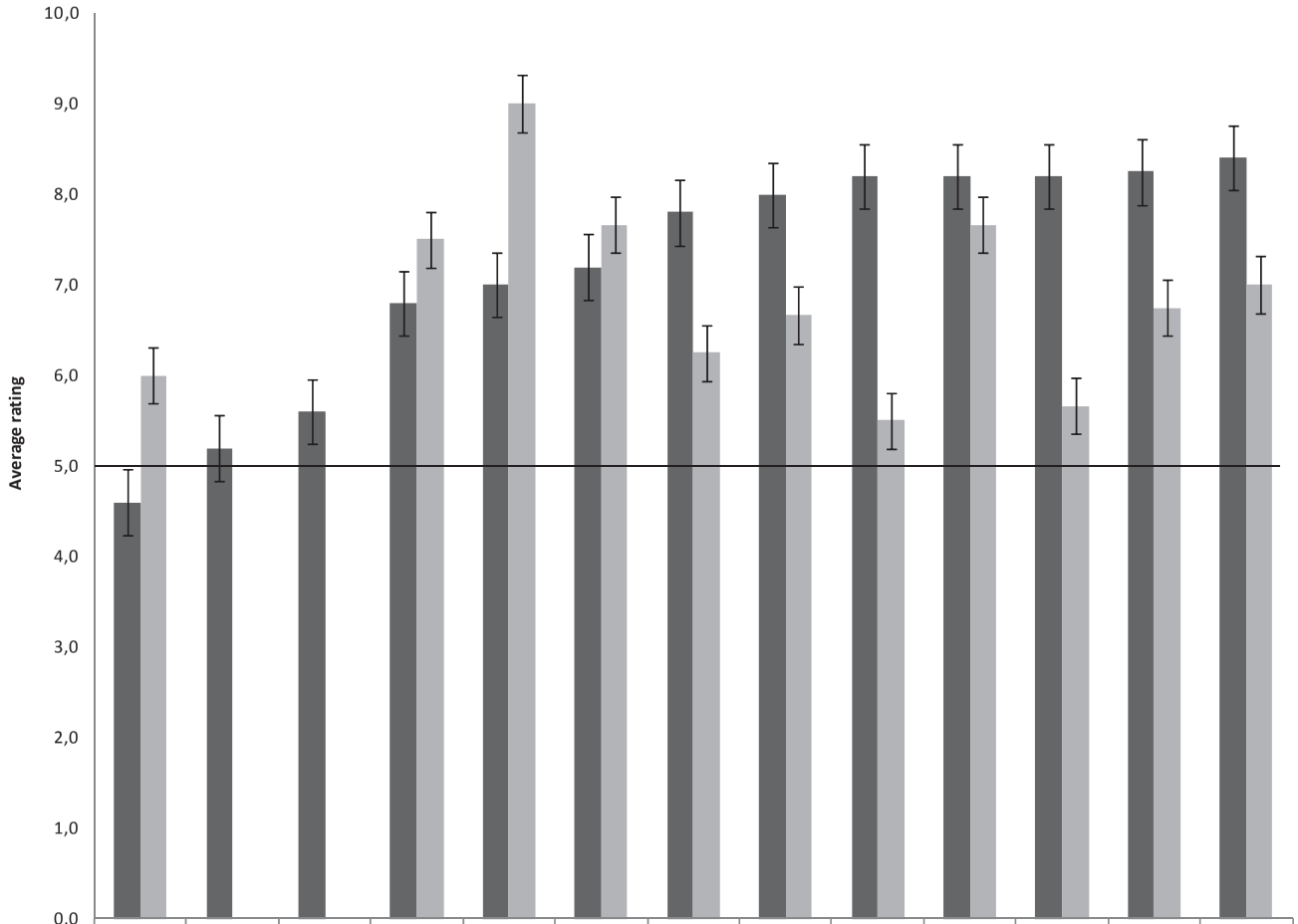
	Creating views using filters	Creating conceptual diagrams individual	Sharing conceptual diagrams	Querying conceptual diagrams using filters	OPP Handbook	Cross-impact method	WoG MA Briefing template	Creation of the common vocabulary	OP Design process	Creation of conceptual diagrams collab.	Common Glossary	OP design tool	Team building
Average usefulness (WoG team)	6,0	6,3	6,4	6,6	7,1	7,3	7,9	8,0	8,4	8,6	8,7	8,7	9,0
Average usefulness (SME observers)		7,1	8,8		7,8	7,6	6,6	5,8	7,8	8,5	7,1	5,1	6,5

Figure 19. Average ratings of SME observers and the WoG team on usefulness-related items by component and sub-component.



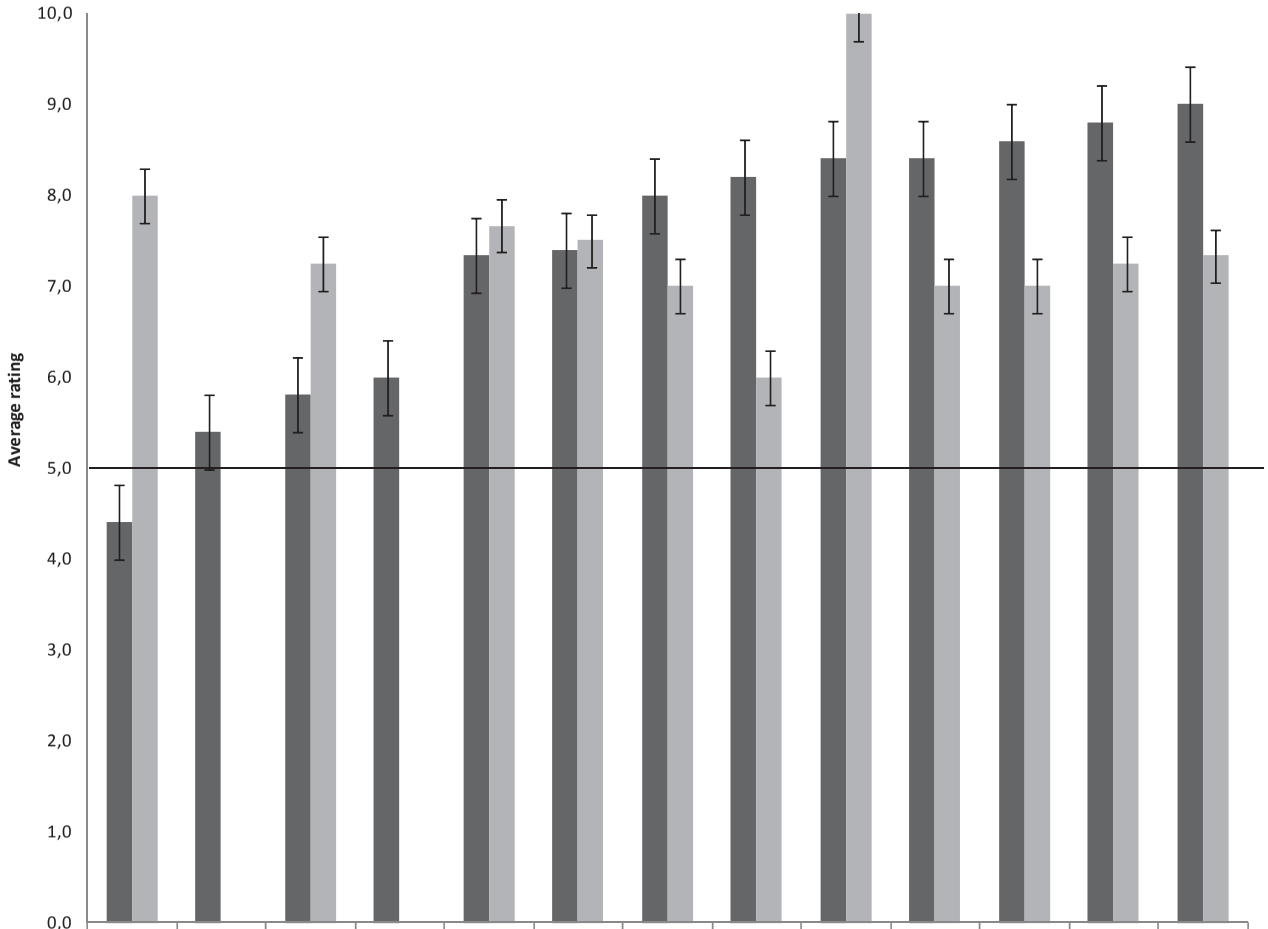
	WoG MA Briefing template	Common Glossary	OP design tool	OPP Handbook	Sharing conceptual diagrams	Team building	Creating views using filters	Querying conceptual diagrams using filters	Creation of the common vocabulary	Creation of conceptual diagrams collab.	Cross-impact method	Creating conceptual diagrams individual	OP Design process
Average time/effort (WoG team)	2,2	2,5	3,3	3,5	3,7	4,1	4,3	4,5	4,5	5,1	5,5	5,7	6,4
Average time/effort (SME observers)	2,4	2,8	3,7	3,2	3,0	2,8			1,5	2,8	6,8	4,0	5,4

Figure 20. Average ratings of observers and the WoG team on time and effort related items by component and sub-component.



	Sharing conceptual diagrams	Querying conceptual diagrams using filters	Creating views using filters	OPP Handbook	Creation of conceptual diagrams collab.	Creating conceptual diagrams individual	WoG MA Briefing template	Cross-impact method	OP design tool	Common Glossary	Team building	OP Design process	Creation of the common vocabulary
Average usability (WoG team)	4,6	5,2	5,6	6,8	7,0	7,2	7,8	8,0	8,2	8,2	8,2	8,3	8,4
Average usability (SME observers)	6,0			7,5	9,0	7,7	6,3	6,7	5,5	7,7	5,7	6,8	7,0

Figure 21. Average ratings of observers and the WoG team on usability related items by component and sub-component. A score below five denotes that the usability of the component is lower than similar tools used in this context. Conversely, a score above five denotes that the usability is higher than similar tools in this context.



	Sharing conceptual diagrams	Querying conceptual diagrams using filters	Creating conceptual diagrams individual	Creating views using filters	Cross-impact method	OPP Handbook	OP design tool	Creation of the common vocabulary	Creation of conceptual diagrams collab.	OP Design process	Common Glossary	WoG MA Briefing template	Team building
Average costs/benefits (WoG team)	4,4	5,4	5,8	6,0	7,3	7,4	8,0	8,2	8,4	8,4	8,6	8,8	9,0
Average costs/benefits (SME observers)	8,0		7,3		7,7	7,5	7,0	6,0	10,0	7,0	7,0	7,3	7,3

Figure 22. Average ratings of observers and the WoG team on cost-benefit related items by component and sub-component. A score below five denotes that costs outweigh the benefits. Conversely, a score above five denotes that the benefits outweigh the costs.

3.2 12om Methodology Multi-criteria Assessment

Using the E-MYRIAD application, it is possible to load a preference model and to create an input file that captures the key metrics collected during LOE #2 in order to perform a series of multi-criteria analyses. E-MYRIAD computes the utility values for each criterion, as well as the satisfaction levels for aggregated variables higher up in the model hierarchy. Results are summarized as percentages and illustrated in a tree with a proportionally filled scale for each node. For the present purposes, any result under 75% is shown in yellow and any result under 50% is shown in red.

E-MYRIAD also performs a sensitivity analysis that systematically tests various input combinations, resulting in an assessment of the differential costs for improving variables. The index labelled "benefit of improving this variable" provides an indication of the relative gain to expect from an eventual improvement in each variable. This index can help select priority areas for future design improvements.

3.2.1 Overall 12om Methodology MYRIAD Assessment

Figure 23 shows the E-MYRIAD results when using the preference model based on the 12om project objectives to perform the overall 12om methodology assessment. Results indicate that the 12om methodology, as used in LOE #2, was 80% successful. Its main strength is in supporting the three dimensions of Collaborative Understanding in a very balanced way. Results in terms of feasibility are relatively good but somewhat lower, mainly due to training time/effort which is seen as demanding and partly due to use time/effort which is seen as somewhat demanding as well.

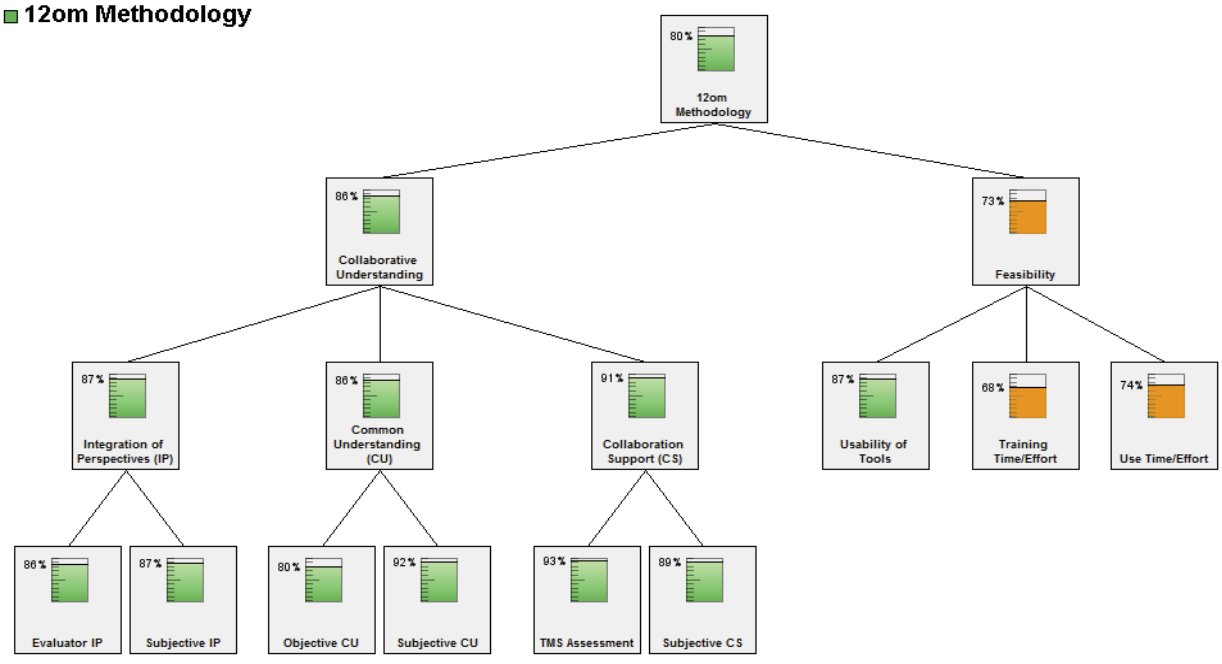


Figure 23: 12om Methodology Multicriteria Assessment

Table 5 shows the "improvement index" for each node in the hierarchy. Since the value is relative to the superordinate variable, the recommended way to interpret these results is to start at the higher levels, looking where there is the most benefit to be obtained, and then tracing down the hierarchy by following the greatest index values to find the priority potential improvement areas.

For instance, collaborative understanding has the greatest improvement index in the higher level of the hierarchy (.50), and under it the Common Understanding node (.47), and under it the Objective Common Understanding metric (.50). In the case of this metric, there was more room for improvement there than anywhere else in the left half of the hierarchy (possibly because this objective metric was less prone to potential biases, e.g., bias to favor novelty). Nonetheless, the Improvement Index is not only sensitive to weaker scores, it takes into account the relationship between factors, which explains that in some cases (e.g., Collaboration Support), there can be actually little or no benefit of improving a variable.

While intuitively, training time/effort could have been identified as the main area for improvement (since this metric technically has the lowest score), E-MYRIAD shows that the greatest value in improving a single metric would come from a change that would impact the Objective Common Understanding metric. Using the above logic makes it straightforward to identify the metric with the greatest improvement value, but it remains unclear how to select the order of the next most valuable metrics to improve. Here we propose using a weighted average to perform the priority ranking. This means calculating the average improvement index along a path in the hierarchy, placing a greater weight on higher level nodes. The lowest nodes (i.e., those shown in the rightmost column of the table) will have a weight of 1, those above (i.e., the middle column) will have a weight of 2, and those above (the leftmost column) will have a weight of three when calculating the average improvement index. For example, the first path shown in Table 5 corresponds to (.50*3 + .36*2 + .50) /6 = .45. This index shows that the first priority should be to improve Objective Common Understanding, and that the second priority should be to improve Subjective Common Understanding.

Table 5. Sensitivity analysis results for the overall 12om methodology

Hierarchy Level and Improvement Index						Priority Metric
Collaborative Understanding	.50					
		Integration of Perspectives (IP)	.36			
				Evaluator IP	.50	***.45***
				Subjective IP	.50	***.44***
		Common Understanding (CU)	.47			
				Objective CU	.50	***.49***
				Subjective CU	.35	***.47***
		Collaboration Support (CS)	.00			
				TMS ratings	.20	***.00***
				Subjective CS	.20	***.00***
Feasibility	.42					
		Usability of Tools	.28			***.36***
		Training Time/Effort	.53			***.46***
		Use Time/Effort	.47			***.44***

3.2.2 Component-specific MYRIAD Assessment

OPP Handbook

Figure 24 shows the E-MYRIAD results for the OPP Handbook component assessment. Results indicate that the OPP Handbook is 71% satisfactory for the purposes of the 12om objectives. Its main weakness is its relatively low impact on taskwork, since it is mainly categorised as reference material. Results in terms of feasibility are relatively good.

Table 6 shows recommendations from the sensitivity analysis. No benefit is expected from improving feasibility since it only improves the overall results when in conjunction with dimensions of support (i.e., when dimensions of support is lower, it constrains the resulting component assessment to its own value). It the present case, improving any of the dimensions of support is deemed equally beneficial.

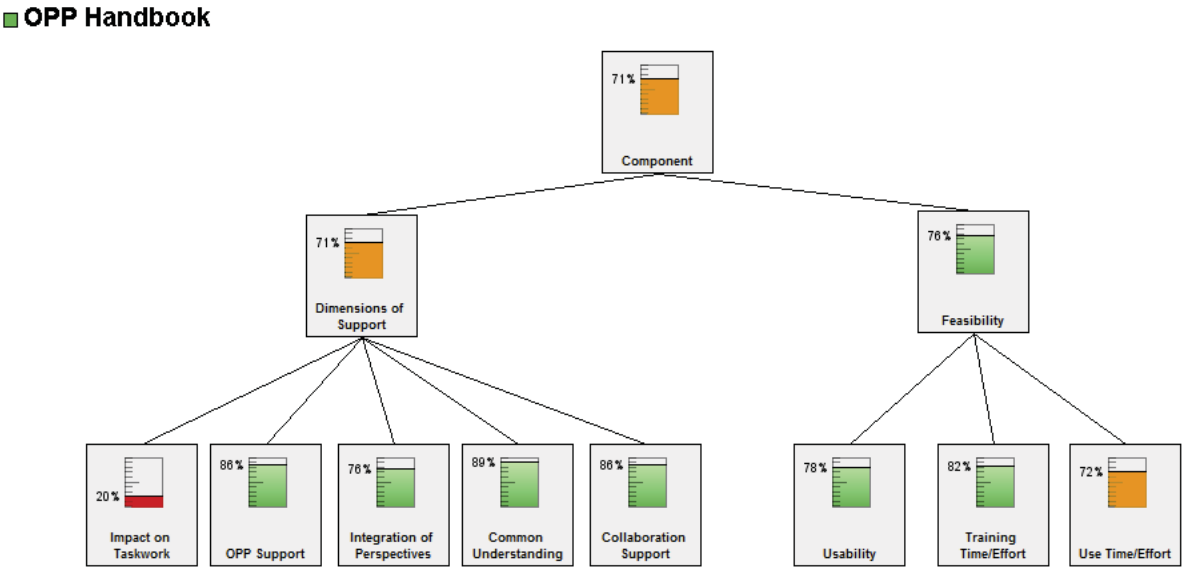


Figure 24: OPP Handbook Multicriteria Assessment

Table 6. Sensitivity analysis results for the OPP Handbook

Hierarchy Level and Improvement Index					Priority Metric
Dimensions of Support	.57				
		Impact on Taskwork	Task-Tool Weight	.50	***.54***
		OPP Support	Ratings (OPP Support)	.50	***.54***
		Integration of Perspectives (IP)	Subjective IP	.50	***.54***
		Common Understanding (CU)	Subjective CU	.50	***.54***
		Collaboration Support (CS)	Subjective CS	.50	***.54***
Feasibility	.00				
			Usability of Tools	.39	***.00***

	Training Time/Effort	.28	***.00***
	Use Time/Effort	.52	***.00***

Team Building

Figure 25 shows the E-MYRIAD results for the Team Building component assessment. Results indicate that this component is 76% satisfactory for the purposes of the 12om objectives. Its main weakness is its relatively low impact on taskwork, since it is mainly categorised as reference material. Otherwise, with the exception of a moderately demanding use time/effort, results are extremely favorable.

Table 7 shows recommendations from the sensitivity analysis. The first priority should be to make design improvements that would increase either OPP support, integration of perspectives, common understanding, or impact on taskwork (although in some cases such as impact on taskwork, an improvement of the design may not be possible due to the component's intrinsic nature, i.e., not being directly related to taskwork).

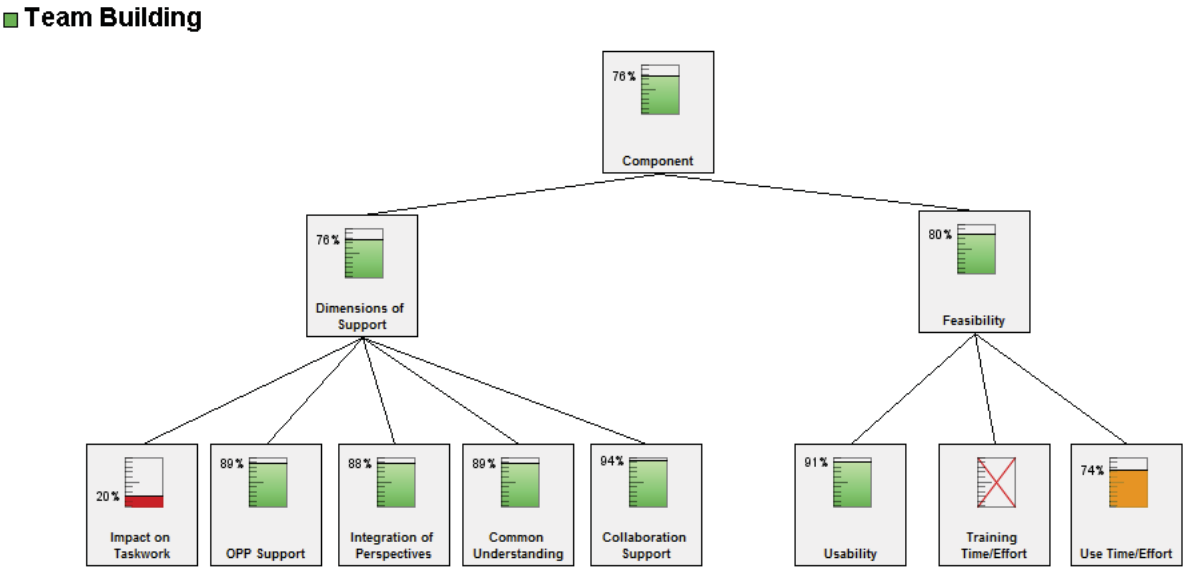


Figure 25: Team Building Procedure Multicriteria Assessment

Table 7. Sensitivity analysis results for the Team Building Procedure

Hierarchy Level and Improvement Index					Priority Metric
Dimensions of Support	.58				
		Impact on Taskwork	Task-Tool Weight	.50	***.55***
		OPP Support	Ratings (OPP Support)	.50	***.55***
		Integration of Perspectives (IP)	Subjective IP	.50	***.55***
		Common Understanding (CU)	Subjective CU	.50	***.55***
		Collaboration Support (CS)	Subjective CS	.12	***.40***
Feasibility	.00				
			Usability of Tools	.39	***.00***
			Training Time/Effort	XX	***XX***

			Use Time/Effort	.53	***.00***

Common Glossary

Figure 26 shows the E-MYRIAD results for the Common Glossary component assessment. Results indicate that the Common Glossary is 76% satisfactory for the purposes of the 12om project. Its sole weakness is its relatively low impact on taskwork, since it is mainly categorised as reference material. Otherwise this component was very well-received.

Table 8 shows recommendations from the sensitivity analysis. The first priority should be to increase either OPP support, Common Understanding, or Collaboration support either OPP support, Common Understanding , or Collaboration support.

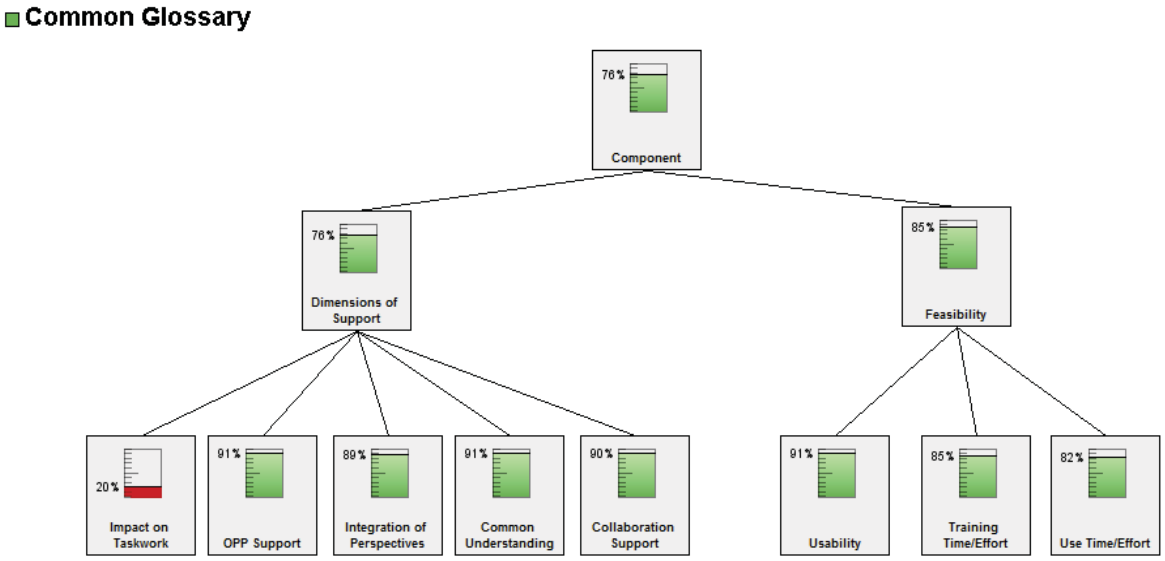


Figure 26: Common Glossary Multi-criteria Assessment

Table 8. Sensitivity analysis results for the Common Glossary

Hierarchy Level and Improvement Index					Priority Metric
Dimensions of Support	.50				
		Impact on Taskwork	Task-Tool Weight	.00	***.00***
		OPP Support	Ratings (OPP Support)	.50	***.50***
		Integration of Perspectives	Subjective IP	.41	***.46***
		Common Understanding	Subjective CU	.50	***.50***
		Collaboration Support	Subjective CS	.50	***.50***
Feasibility	.45				
			Usability of Tools	.28	***.00***
			Training Time/Effort	.49	***.00***
			Use Time/Effort	.53	***.00***

Conceptual Diagrams (Individual)

Figure 27 shows the E-MYRIAD results for the Conceptual Diagrams (Individual) component assessment. Results indicate that this component is 68% satisfactory for the purposes of the 12om objectives. Its main weakness related to feasibility due to the high time and effort required for training and use. Despite these difficulties, this component is deemed highly valuable in terms of its impact on the taskwork to perform.

Table 9 shows recommendations from the sensitivity analysis. The first priority should be to improve any one of dimension of support (other than impact on taskwork which cannot be further improved), closely followed by the training or use time/effort elements.

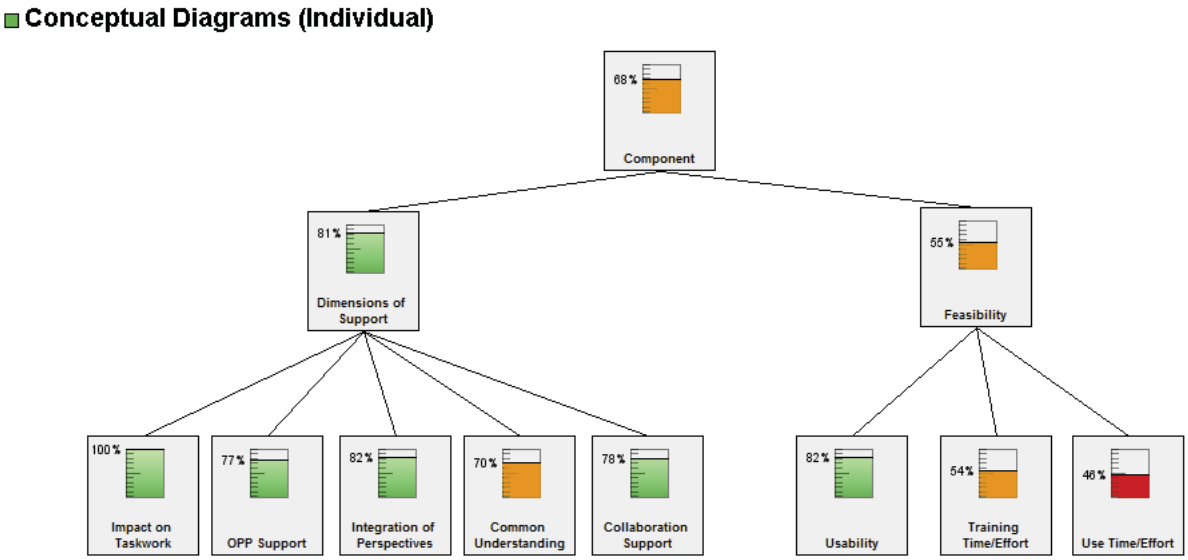


Figure 27: Conceptual Diagrams (Individual) Multicriteria Assessment

Table 9. Sensitivity analysis results for the Conceptual Diagrams (Individual)

Hierarchy Level and Improvement Index					Priority Metric
Dimensions of Support	.50				
		Impact on Taskwork	Task-Tool Weight	.00	***.00***
		OPP Support	Ratings (OPP Support)	.50	***.50***
		Integration of Perspectives (IP)	Subjective IP	.50	***.50***
		Common Understanding (CU)	Subjective CU	.50	***.50***
		Collaboration Support (CS)	Subjective CS	.50	***.50***
Feasibility	.46				
			Usability of Tools	.28	***.39***
			Training Time/Effort	.48	***.47***
			Use Time/Effort	.53	***.49***

Conceptual Diagrams (Collaborative)

Figure 28 shows the E-MYRIAD results for the Conceptual Diagrams (Collaborative) component assessment. Results indicate that this component is 76% satisfactory for the purposes of the 12om objectives. Its main weakness relates to feasibility, specifically its demanding use time/effort. Nonetheless, this component is deemed highly valuable in terms of its impact on the taskwork to perform, and is the component with the highest (94%) satisfaction in terms of *Dimensions of Support*.

Table 10 shows recommendations from the sensitivity analysis. The priority should be to improve feasibility, by improving use time/effort, closely followed by training time/effort.

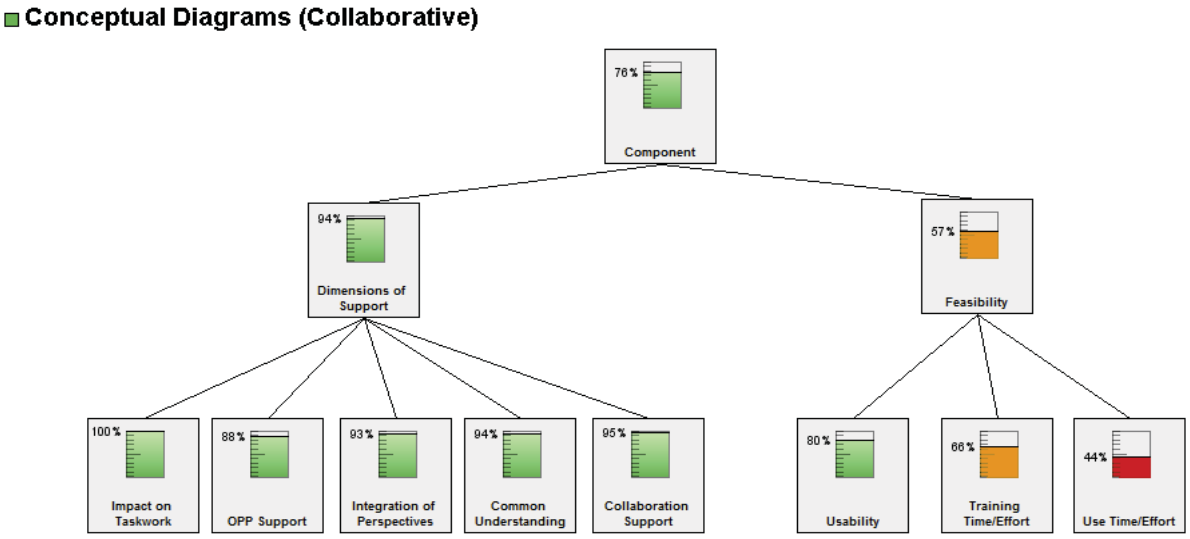


Figure 28: Conceptual Diagrams (Collaborative) Multicriteria Assessment

Table 10. Sensitivity analysis results for the Conceptual Diagrams (Collaborative)

Hierarchy Level and Improvement Index					Priority Metric
Dimensions of Support	.08				
		Impact on Taskwork	Task-Tool Weight	.00	***.00***
		OPP Support	Ratings (OPP Support)	.50	***.24***
		Integration of Perspectives (IP)	Subjective IP	.16	***.11***
		Common Understanding (CU)	Subjective CU	.08	***.08***
		Collaboration Support (CS)	Subjective CS	.00	***.00***
Feasibility	.50				
			Usability of Tools	.28	***.41***
			Training Time/Effort	.45	***.48***
			Use Time/Effort	.55	***.52***

Cross-Impact Method

Figure 29 shows the E-MYRIAD results for the Cross-Impact Method assessment. Results indicate that this component is 63% satisfactory for the purposes of the 12om objectives. Its main weakness is feasibility, specifically in terms of its demanding use time/effort and training time/effort. Nonetheless, this component is deemed highly valuable in terms of its impact on the taskwork to perform and received quite high usability ratings when considering the nature of the analytical method.

Table 11 shows recommendations from the sensitivity analysis. Here, several options for improving the cross-impact method seem equally viable: Use time/effort, OPP Support, Integration of Perspectives, Common Understanding, Collaboration Support, and training time/effort. The choice may therefore go to an area that can be more easily improved by the design team, such as optimising training time/effort using a multimedia tutorial or redesigning part of the process to reduce use time/effort.

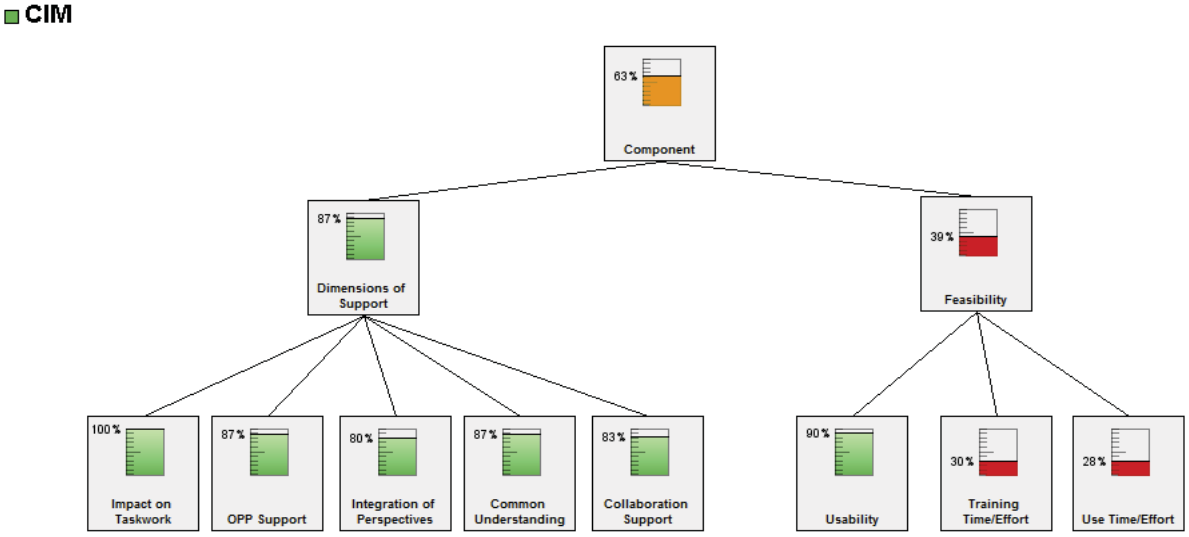


Figure 29: Cross-Impact Method Multi-criteria Assessment

Table 11. Sensitivity analysis results for the Cross-Impact Method

Hierarchy Level and Improvement Index					Priority Metric
Dimensions of Support	.50				
		Impact on Taskwork	Task-Tool Weight	.00	*.00*
		OPP Support	Ratings (OPP Support)	.50	*.50*
		Integration of Perspectives (IP)	Subjective IP	.50	*.50*
		Common Understanding (CU)	Subjective CU	.50	*.50*
		Collaboration Support (CS)	Subjective CS	.50	*.50*
Feasibility	.49				
			Usability of Tools	.28	*.41*
			Training Time/Effort	.50	*.49*

		Use Time/Effort	.51	*.50*

Op Design Tool

Figure 30 shows the E-MYRIAD results for the Op Design Tool assessment. Results indicate that this component is 82% satisfactory for the purposes of the 12om objectives – making it the most successful component in the 12om methodology. Its major strength is that it is well-balanced, yielding the most satisfactory outcome even if this component is not ranked first in terms of feasibility or dimensions of support.

Table 12 shows recommendations from the sensitivity analysis. The most valuable areas to improve are Impact on taskwork, Integration of Perspectives, Common Understanding, and Collaboration Support.

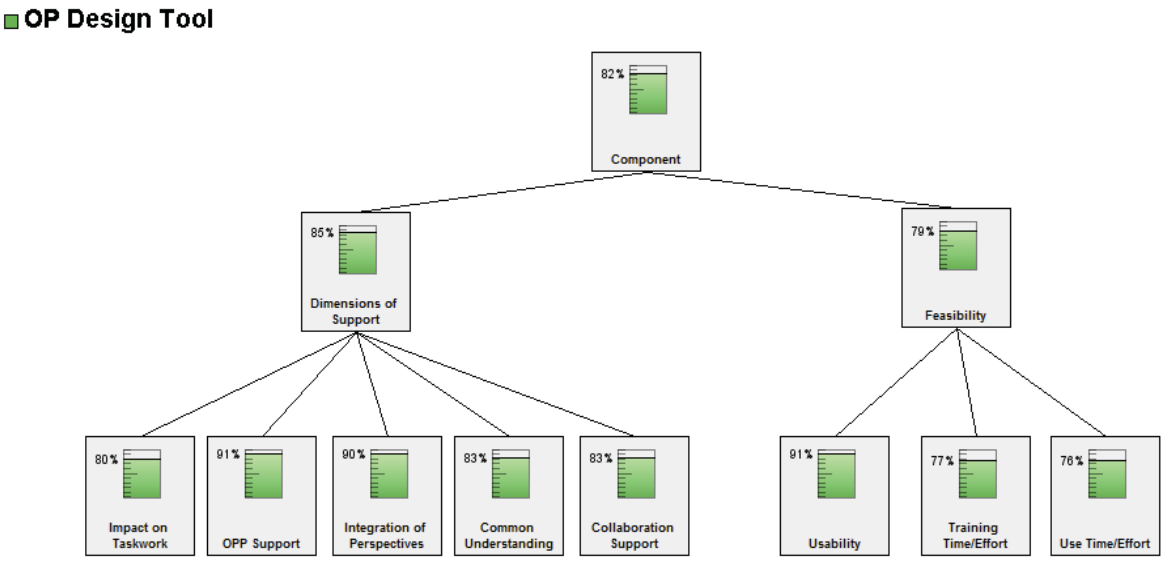


Figure 30: Op Design Tool Multicriteria Assessment

Table 12. Sensitivity analysis results for the Op Design Tool

Hierarchy Level and Improvement Index					Priority Metric
Dimensions of Support	.50				
		Impact on Taskwork	Task-Tool Weight	.50	***.50***
		OPP Support	Ratings (OPP Support)	.43	***.47***
		Integration of Perspectives (IP)	Subjective IP	.50	***.50***
		Common Understanding (CU)	Subjective CU	.50	***.50***
		Collaboration Support (CS)	Subjective CS	.50	***.50***
Feasibility	.35				
			Usability of Tools	.28	***.32***
			Training Time/Effort	.48	***.40***
			Use Time/Effort	.51	***.41***

WoG MA Brief Template

Figure 31 shows the E-MYRIAD results for the WoG MA Brief Template assessment. Results indicate that this component is 70% satisfactory for the purposes of the 12om objectives. Its sole weakness is its relatively low impact on taskwork, since it is mainly categorised as reference material. This component has the greatest feasibility amongst the different components considered here.

Table 13 shows recommendations from the sensitivity analysis. The MYRIAD model shows that there is no point in improving feasibility here, and that all five dimensions of support are equally good areas for improvement here.

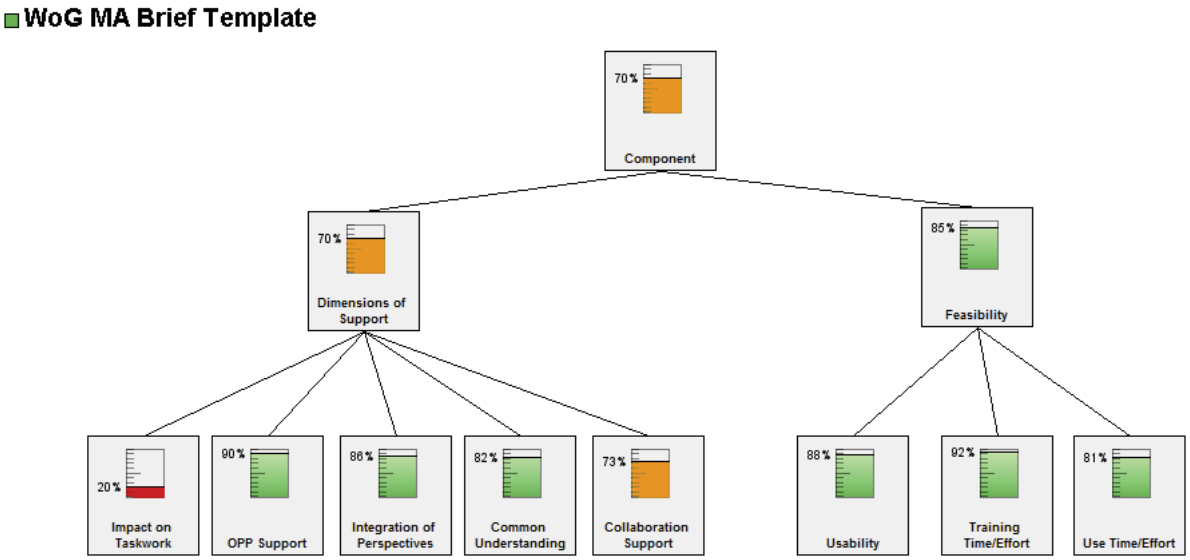


Figure 31: WoG MA Brief Template Multi-criteria Assessment

Table 13. Sensitivity analysis results for the WoG MA Brief Template

Hierarchy Level and Improvement Index					Priority Metric
Dimensions of Support	.63				
		Impact on Taskwork	Task-Tool Weight	.50	***.58***
		OPP Support	Ratings (OPP Support)	.50	***.58***
		Integration of Perspectives (IP)	Subjective IP	.50	***.58***
		Common Understanding (CU)	Subjective CU	.50	***.58***
		Collaboration Support (CS)	Subjective CS	.50	***.58***
Feasibility	.00				
			Usability of Tools	.43	***.00***
			Training Time/Effort	.28	***.00***
			Use Time/Effort	.55	***.00***

Sharing Conceptual Diagrams

Figure 32 shows the E-MYRIAD results for Sharing Conceptual Diagrams. Results indicate that this sub-component is 69% satisfactory for the purposes of the 12om objectives. Its main weakness is its lower usability.

Table 13Table 14 shows recommendations from the sensitivity analysis. The MYRIAD model shows that four equally good areas for improvement are impact on taskwork, integration of perspectives, common understanding, and collaboration support.

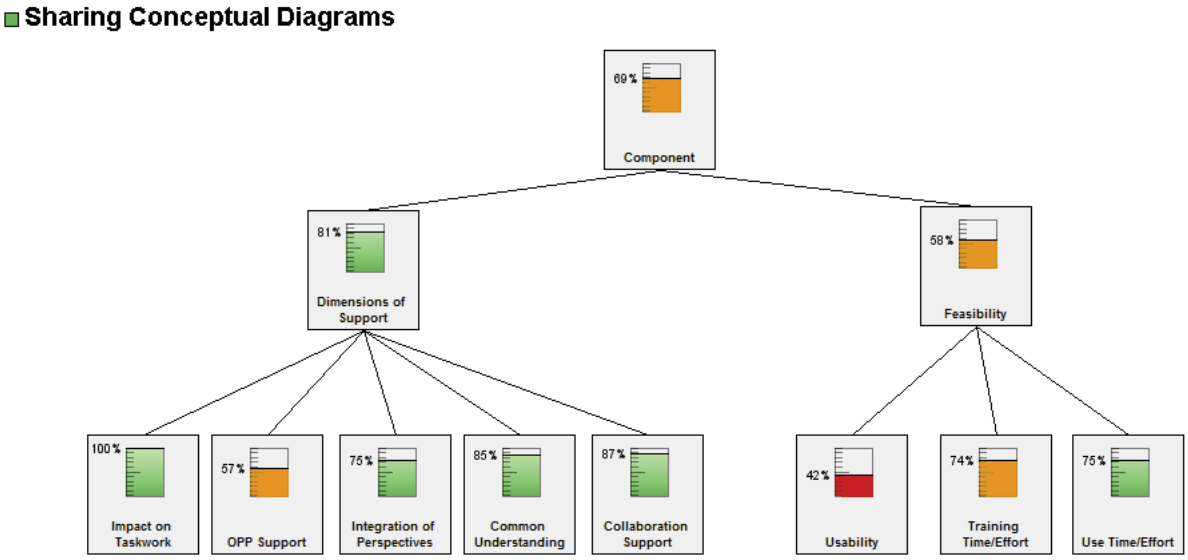


Figure 32: Sharing Conceptual Diagrams Multi-criteria Assessment

Table 14. Sensitivity analysis results for Sharing Conceptual Diagrams

Hierarchy Level and Improvement Index					Priority Metric
Dimensions of Support	.50				
		Impact on Taskwork	Task-Tool Weight	.00	***.00***
		OPP Support	Ratings (OPP Support)	.50	***.50***
		Integration of Perspectives (IP)	Subjective IP	.50	***.50***
		Common Understanding (CU)	Subjective CU	.50	***.50***
		Collaboration Support (CS)	Subjective CS	.50	***.50***
Feasibility	.44				
			Usability of Tools	.56	***.49***
			Training Time/Effort	.31	***.39***
			Use Time/Effort	.28	***.38***

Creation of Common Vocabulary

Figure 33 shows the E-MYRIAD results for the Creation of a Common Vocabulary. Results indicate that this sub-component is 81% satisfactory for the purposes of the 12om objectives. Its main weakness is use time/effort.

Table 15 shows recommendations from the sensitivity analysis. Three equally good priority areas for improvement are OPP support, common understanding, and collaboration support.

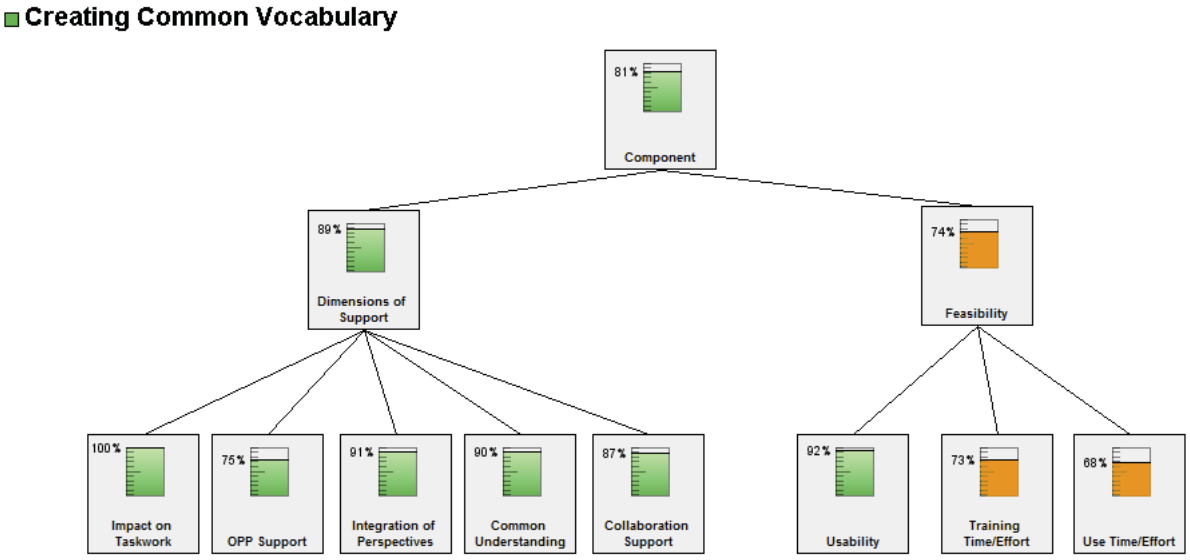


Figure 33: Creation of Common Vocabulary Multi-criteria Assessment

Table 15. Sensitivity analysis results for the Creation of Common Vocabulary

Hierarchy Level and Improvement Index					Priority Metric
Dimensions of Support	.50				
		Impact on Taskwork	Task-Tool Weight	.00	***.00***
		OPP Support	Ratings (OPP Support)	.50	***.50***
		Integration of Perspectives (IP)	Subjective IP	.41	***.46***
		Common Understanding (CU)	Subjective CU	.50	***.50***
		Collaboration Support (CS)	Subjective CS	.50	***.50***
Feasibility	.45				
			Usability of Tools	.28	***.38***
			Training Time/Effort	.49	***.47***
			Use Time/Effort	.53	***.48***

Creating Views Using Filters

Figure 34 shows the E-MYRIAD results for Creating Views Using Filters. Results indicate that this component is 68% satisfactory for the purposes of the 12om objectives. Its main weakness is its relatively low OPP support.

Table 13Table 16 shows recommendations from the sensitivity analysis. The MYRIAD model shows that four equally good priority areas for improvement are OPP support, integration of perspectives, common understanding, and collaboration support.

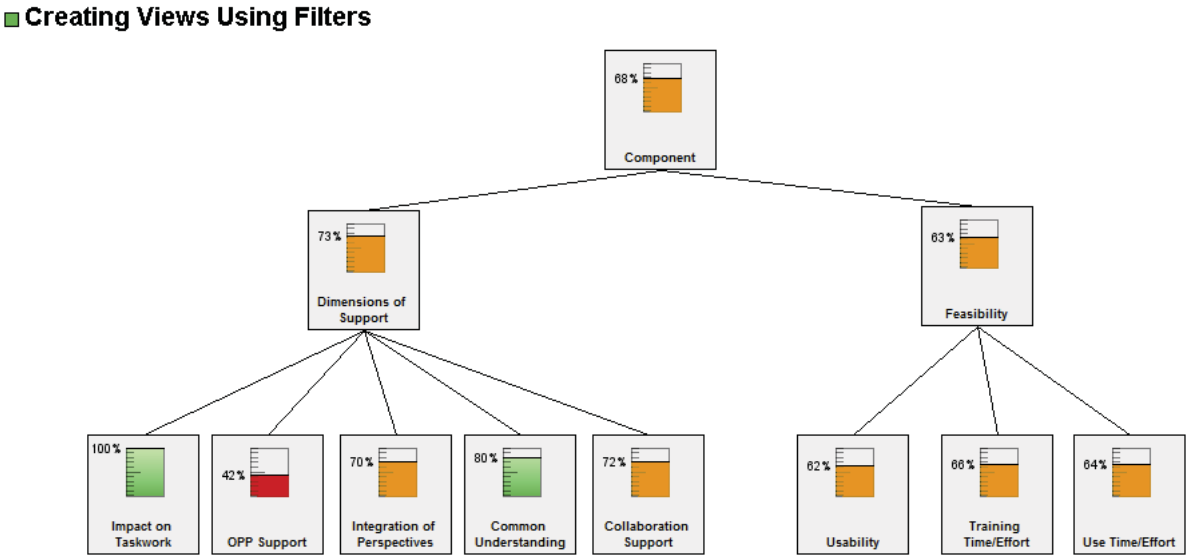


Figure 34: Creating Views Using Filters Multi-criteria Assessment

Table 16. Sensitivity analysis results for Creating Views Using Filters

Hierarchy Level and Improvement Index					Priority Metric
Dimensions of Support	.50				
		Impact on Taskwork	Task-Tool Weight	.00	***.00***
		OPP Support	Ratings (OPP Support)	.50	***.50***
		Integration of Perspectives (IP)	Subjective IP	.50	***.50***
		Common Understanding (CU)	Subjective CU	.50	***.50***
		Collaboration Support (CS)	Subjective CS	.50	***.50***
Feasibility	.32				
			Usability of Tools	.40	***.35***
			Training Time/Effort	.28	***.30***
			Use Time/Effort	.32	***.32***

Querying Diagrams Using Filters

Figure 35 shows the E-MYRIAD results for Querying Diagrams Using Filters. Results indicate that this component is 68% satisfactory for the purposes of the 12om objectives. Its key weakness is its relatively low OPP support, usability, and use time/effort.

Table 17 shows recommendations from the sensitivity analysis. The MYRIAD model identifies four priority areas for improvement: OPP support, integration of perspectives, common understanding, and collaboration support.

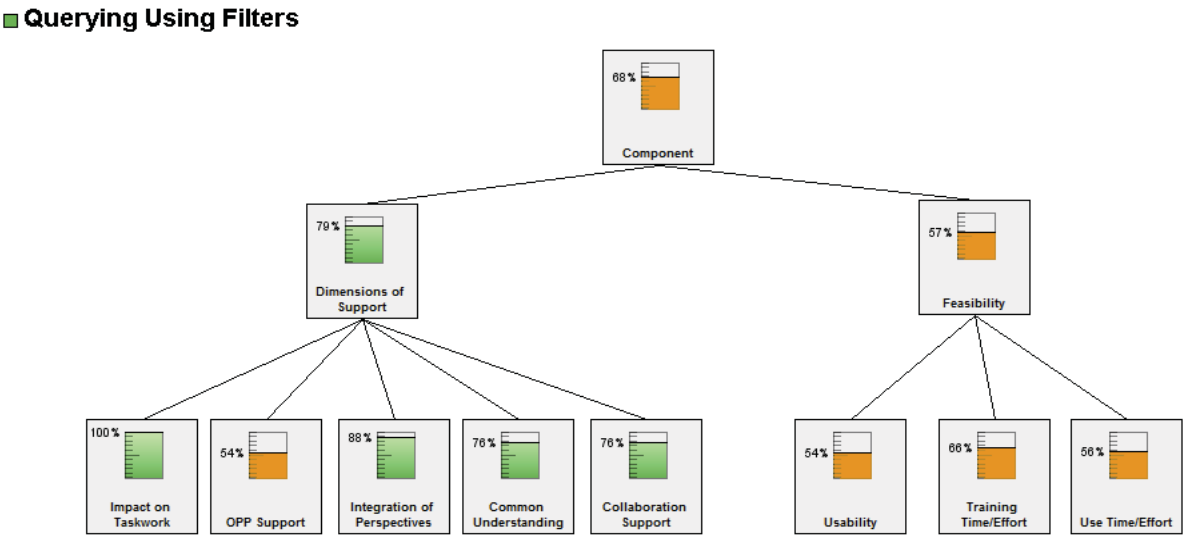


Figure 35: Querying Diagrams Using Filters Multi-criteria Assessment

Table 17. Sensitivity analysis results for Querying Diagrams Using Filters

Hierarchy Level and Improvement Index					Priority Metric
Dimensions of Support	.50				
		Impact on Taskwork	Task-Tool Weight	.00	***.00***
		OPP Support	Ratings (OPP Support)	.50	***.50***
		Integration of Perspectives (IP)	Subjective IP	.50	***.50***
		Common Understanding (CU)	Subjective CU	.50	***.50***
		Collaboration Support (CS)	Subjective CS	.50	***.50***
Feasibility	.43				
			Usability of Tools	.45	***.44***
			Training Time/Effort	.28	***.37***
			Use Time/Effort	.40	***.42***

3.3 Performance Analysis

Performance analysis is distinct from the 12om methodology assessment presented so far because it concerns goal attainment of the WoG planning team in terms of mission success rather than the evaluation of the methodology components. Performance was assessed through the use of measures of performance (MoP) and measures of effectiveness (MoE). Combined, they provide a fairly good portrayal of the WoG planning team's performance.

3.3.1 Measures of Performance

NASA-TLX assesses performance with five questionnaire items: (1) mental demand, (2) temporal demand, (3) effort, (4) frustration, and (5) performance. The scale used during this study ranged from 1 to 10, with a middle point at 5.5. Apart from performance, for which you want to have the highest score possible, the target value would be a middle point. At this point, the workload is believed to be at an optimal point given a non-linear "inverse U-shaped" function between workload and performance. Average ratings are represented by scale in Figure 36.

Overall, NASA-TLX ratings suggest very good performance from the WoG team as mental demand, temporal demand, effort and frustration are all around 5.5, and performance is relatively high (i.e., 7.8). Moreover, the relatively low variability across team members (see error bars on Figure 36) suggests that all members were generally around the optimal workload level (i.e., there was a good workload balance throughout the team).

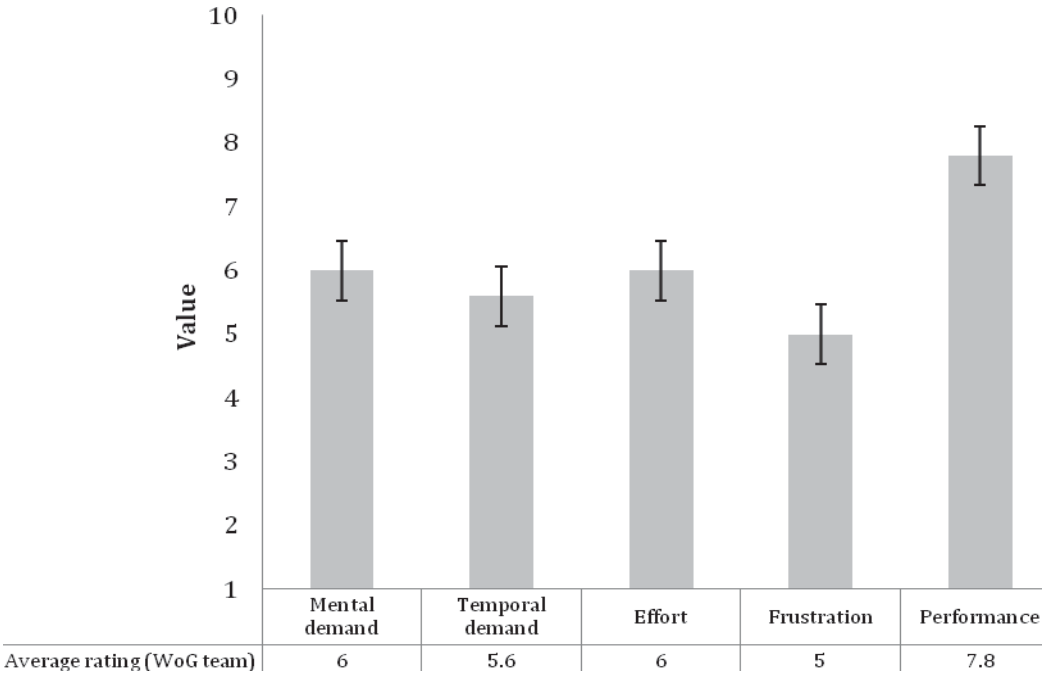


Figure 36. NASA-TLX average ratings of the WoG team members by dimension. Error bars show standard error.

The Mission Awareness Rating Scale (MARS) is a self-rating assessment technique designed specifically for use in the assessment of situation awareness (SA) during a military exercise, and provides information on the participants' capacity to acquire and maintain SA. It comprises two separate sets of questions – ability to acquire SA and difficulty to maintain SA - based on Endsley's (1995) three-level model of SA (i.e., perception, comprehension and projection).

Figure 37 represents average ratings of the perceived ability to achieve each level of SA and of the awareness of how to achieve mission goal. Overall, results suggest that there were no particular difficulties associated with gaining SA during the LOE #2, independently from the level of SA sought.

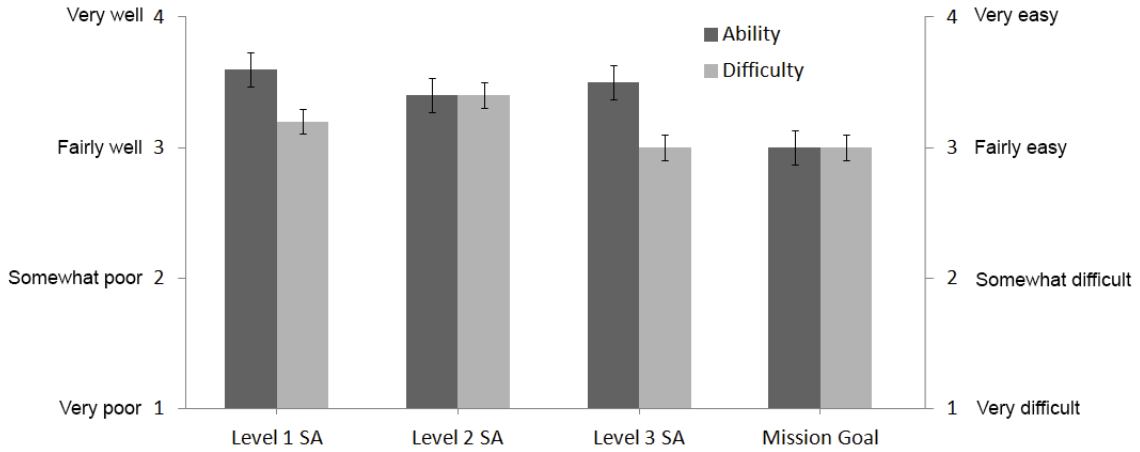


Figure 37. Ability to acquire/difficulty to maintain SA by level of SA. Error bars show standard error.

The process evaluation questionnaire aims to assess the performance of the integrated planning team in terms of process rather than outputs. The scale ranged from 1 (not at all) to 5 (very much) with 3 being average. Figure 38 shows average ratings of the SME evaluators (n = 3, the ratings from 1 SME were missing). Results show that the lowest average ratings were 2. Three items were rated 2. They are reported below and comments pertaining to these items are shown when relevant:

- How effectively did the team examine force capability and groupings?
  - *Not really articulated, but less critical in this case.*
- How effectively did the team identify key strengths and weaknesses?
  - *Not well articulated, probably due in part by the lack of slide in the MA brief format.*
- How thoroughly did the team challenge its own assumptions?
  - *The assumptions were developed, but most did not meet the requirement of "necessity" although most were "reasonable".*

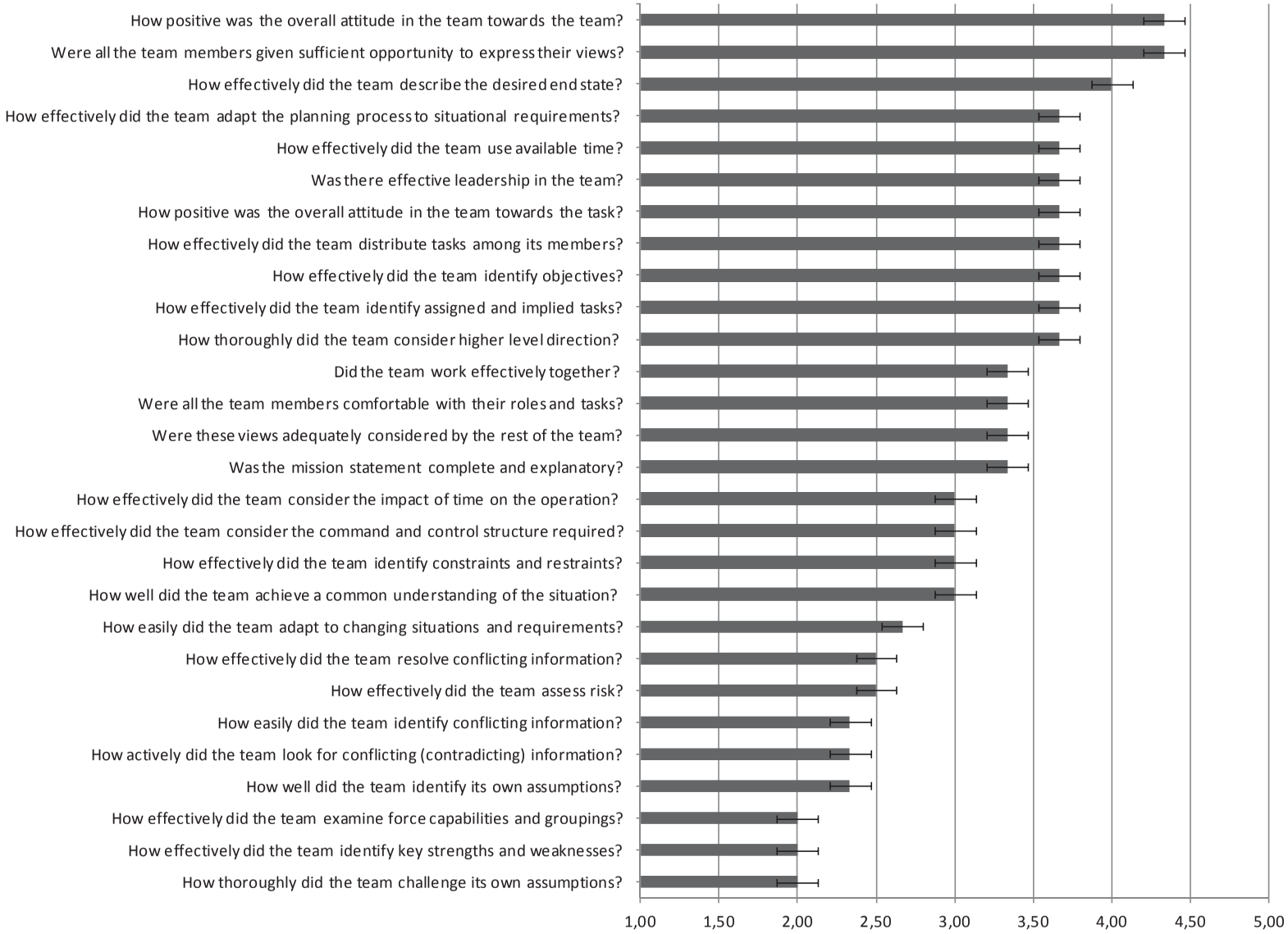


Figure 38: Average ratings of the SME evaluators by item of the process evaluation questionnaire.

3.3.2 Measure of Effectiveness

The measure of effectiveness selected for this LOE was based on an evaluation of the SME observers of the MA brief product created by the WoG planning team. The evaluation was composed of six items rated on a scale from 1 to 5 (1 = "Does not meet expectations", 3 = "Meets expectations", and 5 = "Exceeds expectations"). An overall score (overall effectiveness) representing the average of the six items was also calculated. Results of the evaluation are represented in Figure 39.

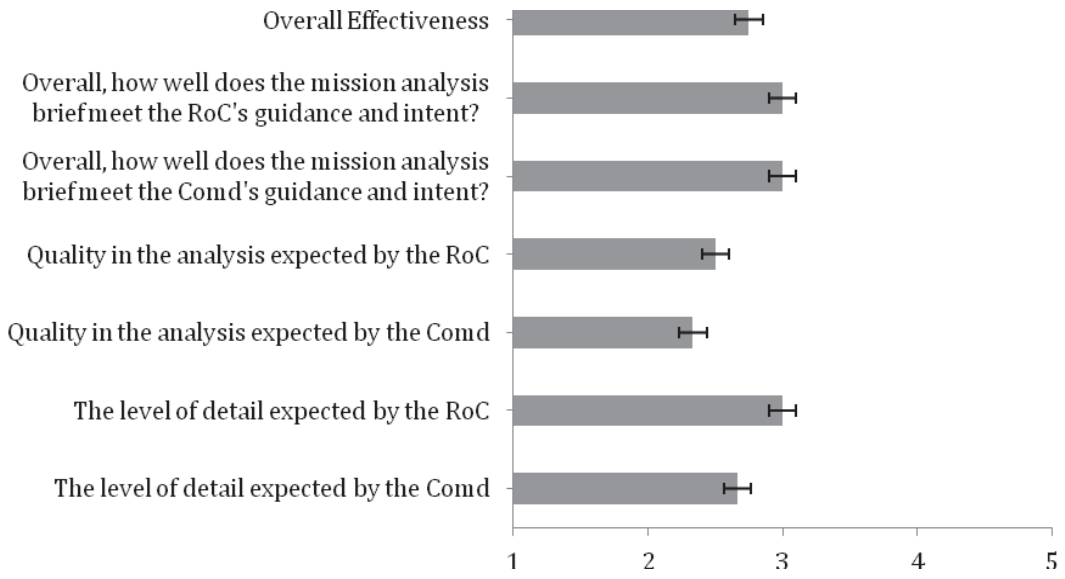


Figure 39. Evaluation of the performance of the WoG planning team by the SME observers. Mean ratings shown. Error bars represent standard error.

All ratings were around a value of 3 (i.e., meet expectations). This is especially good since a lot of the time for conducting mission analysis was actually used for training on the 12om components or filling questionnaires. On the other hand, the ratings show place for improvement as the team never exceeded expectations of the raters. In other words, they performed relatively well, but they could have been much better.

3.4 Common Understanding

The WoG planning team members were asked to list the top 10 influencing scenario-related factors that impacted either positively or negatively on mission success. An integrated list was generated, containing all non-overlapping factors listed individually. Each WoG team member was then asked to rate the factors listed in the integrated list (their own factors and the factors listed by other team members) relatively to their potential impact on mission success. The scale was ranging from -3 (strong unfavourable impact on mission success) to +3 (strong favourable impact on mission success), with 0 being a neutral point (no impact on mission success). Answers were then correlated across WoG team members in order to assess how much they agree on the impact of different

factors relatively to their goal, or common understanding of the mission. Correlation coefficients are reported in Table 18.

Table 18. Correlations between the "mental models" of the WoG planning team members.

	J5	J5Ops	J52	J5Gov	J5Dev	Average
J5	-	0.52	0.85	0.81	0.73	0.73
J5Ops	0.52	-	0.67	0.66	0.57	0.61
J52	0.85	0.67	-	0.89	0.71	0.78
J5Gov	0.81	0.66	0.89	-	0.75	0.78
J5Dev	0.73	0.57	0.71	0.75	-	0.69
					Average:	0.72

Note. The mental model was estimated by asking the participants to rate the impact of several scenario-related factors on mission success.

Results show that average correlation is strong (i.e., $r = .72$) which suggests that common understanding of the mission factors is good. Interestingly, results show that there is no major discrepancy between civilian and non-civilian agencies, which is an indication of integration of perspectives into the common understanding of the WoG team. Another interesting finding is the lower correlation for the J5Ops ($r = .61$) compared to the other team members' correlation coefficients. This may be explained by the fact that this particular individual did not participate in the cross-impact method exercise, which is directly relevant in formalizing the impact of mission factors relatively to mission success. The degree of overlap between individual factors was also assessed. Theoretically, the team could have had identified up to 50 different factors (i.e., when merging their individual list of 10 most important factors). However, the merged list resulted in 31 factors. This result indicates that there was some level of overlap between the team members, but that the lists were also relatively complementary.

Inter-mission achievement factors were also analysed. On average, J5 identified 1.25/3 factors important to the other members of the planning team. J52 and J5Dev identified 2.25/3, J5Gov identified 2.5/3, and J5Ops identified 2.75/3. Considered together, team members identified on average 2.2/3 factors important to the other members. This result shows that team members were generally aware of what was important to the other ones.

3.5 Functional Gap Analysis

A functional gap analysis was performed with the data collected through the task-to-tool mapping exercise carried out at the end of the LOE #2. During this exercise, WoG team members individually had to link 12om components to the tasks/sub-tasks associated with initiation, orientation and COA development phases of the OPP. They were free to add any sub-task or tools that they felt were missing from the template that was given to them. Afterward, they were asked to rate the level of support (between 0 and 3, where 0 is very low support and 3 is very high support) of each component. Once completed, the WoG team members agreed on a collaborative mapping by discussing their individual mapping. Figure 40 shows the final collaborative mapping. This map enabled a functional gap analysis, which is the identification of any gaps in the tasks that would require support but that are

not supported, or conversely, the identification of any component that does not provide any support to the targeted tasks of the OPP.

Is there a task of the OPP that is not supported by any component of the 12om methodology?

The simple answer to this question is no. Every task or sub-task of the OPP from the initiation to the COA development is supported by at least two components of the 12om methodology. However, a more detailed analysis suggests a more complex answer to this question. Table 19 summarizes support of components to OPP tasks and sub-tasks and 12om objectives. Each line represents a component, and each column represents a task, sub-task, or objective. Every time that the WoG team linked a component to a task (meaning that the component supports this particular task) the matrix indicates a "1".

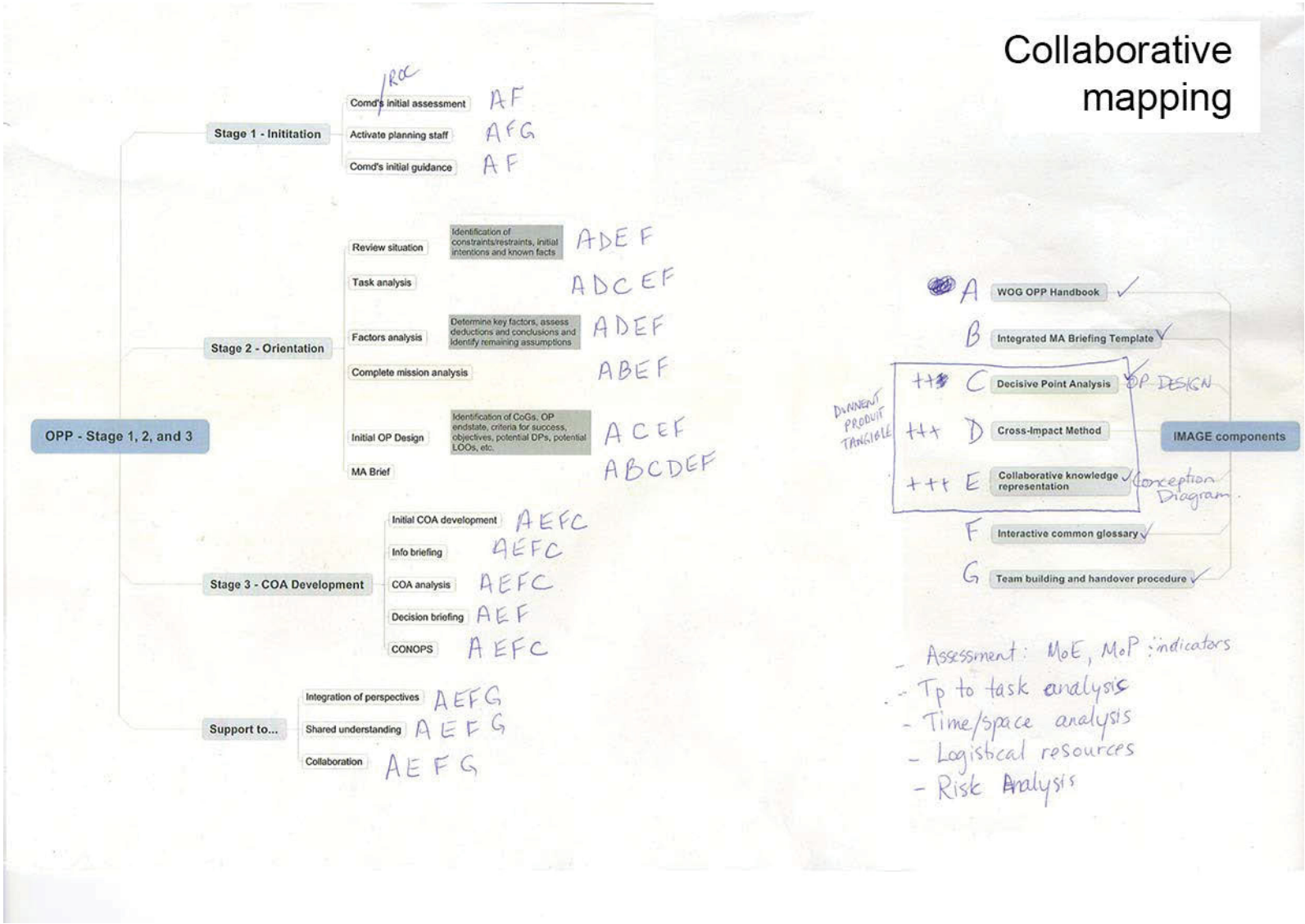


Figure 40. Collaborative mapping result.

Support provided by each component. The sum of "1s" on a line gives a sense of the flexibility of application of the component across the three first stages of OPP and in supporting the more general 12om objectives. A high sum suggests that the component is useful for a wide variety of tasks/objectives (i.e., flexible), and vice versa for a low sum (i.e., not flexible). This sum was also weighed according to the impact on taskwork afforded to the component by the WoG team members. Impact on taskwork was also reported in the multi-criteria presented above. The impact, ranging from 0 to 4, is represented between parentheses in Table 19. The range corresponds to the 0 - +++ range that the participants used during the exercise. According to the participants, the OP Design tool was weighted "3", and cross-impact method and collaborative knowledge representation were weighted "4". For these components, the value of a support (i.e., 1) was multiplied by the weight of the impact before summation. The weighted sum consequently represents both flexibility and impact of the component. We refer to the weighted sum as the level of support provided by a component.

Table 19. Summary of support to OPP and 12om objectives by methodology component.

	Initiation			Orientation						COA development					Objectives				
	Comd's/ROC initial assessment	Activate planning staff	Comd's initial guidance	Review siuation	Task analysis	Factors analysis	Complete mission analysis	Initial OP Design	MA Brief	Initial COA development	Info briefing	COA analysis	Decision briefing	CONOPS	Integration of perspectives	Shared understanding	Collaboration	Sum	Weighed sum
WoG OPP handbook	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	17	17
Integrated MA briefing template							1(1)		1(1)									2	2
OP Design tool					1(3)			1(3)	1(3)	1(3)	1(3)	1(3)		1(3)				7	21
Cross-impact method				1(4)	1(4)	1(4)			1(4)									4	16
Collaborative KR				1(4)	1(4)	1(4)	1(4)	1(4)	1(4)	1(4)	1(4)	1(4)	1(4)	1(4)	1(4)	1(4)	1(4)	14	56
Interactive common glossary	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	17	17
Team building and handover		1(1)													1(1)	1(1)	1(1)	4	4
Sum	2	3	2	4	5	4	4	4	6	4	4	4	3	4	4	4	4	65	
Weighed sum	2	3	2	10	13	10	6	7	14	9	9	9	6	9	7	7	7		133

Results show that collaborative knowledge representation was the component that provided the highest level of support to the OPP tasks and 12om objectives as a whole (i.e., weighed sum of 56). This is mainly due by the facts that the component was very flexible (as denoted by the un-weighed sum of 14 out of 17) and was qualified with a strong impact on task work (as denoted by the weight of "4"). In other words, this component is at the same time general in its context of application and specialized in its features, allowing for high level support.

Results also suggest that two components, interactive common glossary and WoG OPP handbook, do not provide a high level of support, but are however useful throughout the whole process, henceforth granting them an overall respectable level of support. In other

words, these components are very flexible in their context of application, but are not specialized enough to grant specific, high level support to any particular tasks.

OP Design tool and cross-impact method also achieve respectable levels of overall support, mainly because of their impact, rather than relying on the flexibility. Indeed, they do not provide support to many tasks and sub-tasks of the OPP; however, they are specialized enough so that the support provided is important.

Finally, some of the components, Integrated MA briefing template and team building and handover procedure are neither flexible, nor they have a strong impact on task work. Consequently, although they still are relevant in the context of OPP, they provide a limited level of support. On a side note, team building and handover procedure mostly support 12om objectives, but not OPP tasks specifically. This suggest that this method is not focused on supporting OPP per se, but more on underlying cognitive/social functions that are at play during OPP.

Level of support to task. The sum of the "1s" in a column represents the level of support provided to the task corresponding to this column. A high sum suggests that the task is supported by a wide variety of components, and vice versa for a low sum. This sum was also weighted according to the level of support afforded to the component by the WoG team members.

Results show that tasks underlying Orientation phase of the OPP are the most supported (both in terms of weighed and un-weighed support) by the 12om methodology components. On the other side, the average support granted by the 12om methodology components to the tasks of the initiation phase is somewhat lower. This is not dramatic though, as the main focus of support by the 12om methodology is on the orientation and COA development stages of the OPP.

Are there other tools or tasks that would be relevant in similar planning contexts?

The functional gap analysis also allows the identification of potential tasks and/or tools that are missing from the original template given to the participants. In the context of the current study, WoG team members suggested the following tasks that could benefit to the planning process in general:

- Assessment: MoE, MoP indicators;
- Time/space analysis;
- Logistical resources; and
- Risk analysis.

Some of these tasks are however already included within the OPP, but were not represented in the initial template provided to the participants because they were less relevant in the context of the specific scenario used during this study (i.e., they were implicitly included in the larger categories shown).

3.6 Team Dynamics

Team dynamics refer to the analyses of the WoG team members' interactions. Social network analysis informs about the relative importance of each team member during the LOE. The relative importance is determined by taking into account the level of involvement of each team member in the discussions. Content analysis describes the nature of the communications at different times during the experiment. It allows identifying transition phases in the planning process and relative importance of the different lines of operations that usually describes integrated planning efforts (i.e., security, governance, and development). Finally, transactive memory systems analysis characterizes the team's functional potential in terms of coordination and credibility.

3.6.1 Social Network Analysis

Social network analysis consisted, in the context of the current study, of three metrics – emission degree, reception degree, and sociometric status (e.g. Benta, 2005) – and one visual representation. The analyses were conducted on the communication interaction data (i.e., frequency of communications between agents) that was collected during the experiment.

Emission degree of a team member is the sum of all the frequencies of communications originating from that member. The reception degree of a team member is the sum of all frequencies of communications directed to that same team member. Sociometric status is a measure of 'how busy' team member is relative to the overall number of members in the team. In practical terms, sociometric status gives an indication of the relative prominence an individual has as a communicator with others in the team. The sociometric status of a node is the sum of its reception and emission degrees, relative to the number of all other nodes in the network:

$$status(i) = \frac{1}{g-1} \sum_{j=1}^{g} (x_{ji} + x_{ij})$$

where g is the number of node, i is the index of the current node and x_{ji} are the edge values from node j to node i.

All the data necessary for the calculation of these metrics are shown in Figure 41.

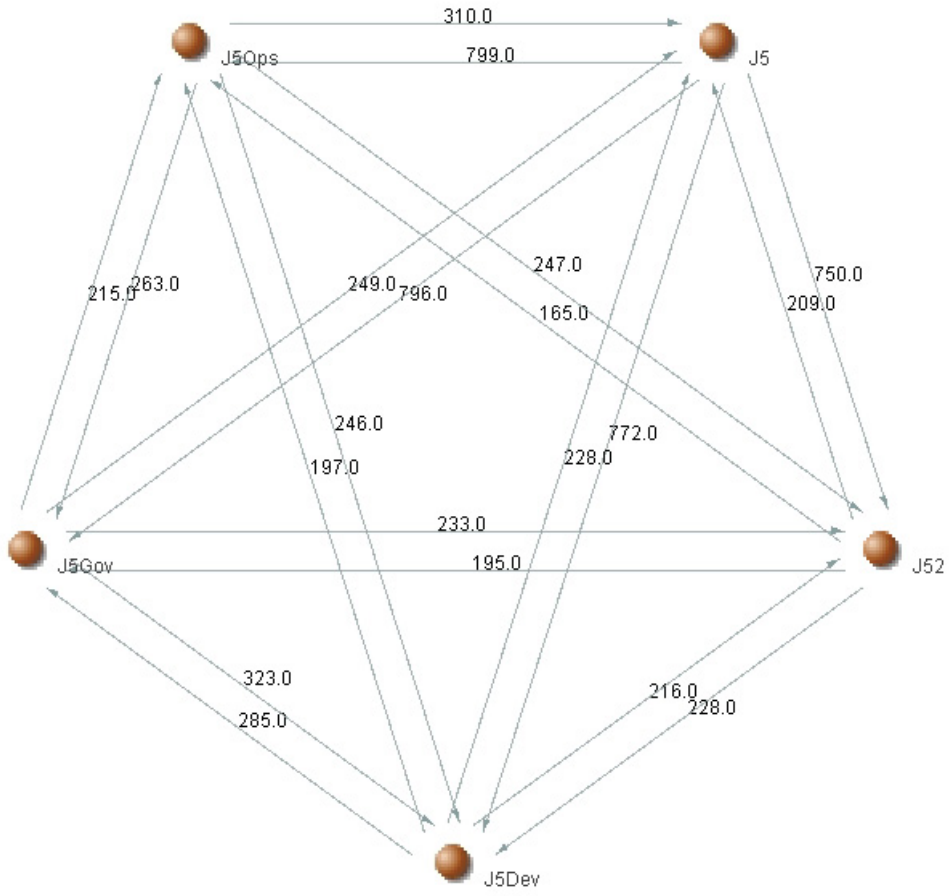


Figure 41. Social network visual representation. Darker links signify high number of communications. Links are not bidirectional as a node can have a different amount of incoming/outgoing communications. Numbers close to each node indicate the amount of outgoing communications from that node.

Results

The values of emission degree, reception degree and sociometric status are reported by role in Table 20. Results concerning emission degree show that the J5 talks much more than the other individuals in the team. This is an indication that the J5 assumed his role of team leader during the exercise. Results also show that J52 was the least talkative individual of the team, which is not alarming since there were three individual out of five from the military. Moreover, considering the role of a J52 within an integrated planning team, you would expect individuals with this role to talk less (and focus on carrying analyses). The other roles all had similar emission degrees, which is a sign that nobody was left out or took over all discussions.

Table 20. Emission degree, reception degree and sociometric status by role in the WoG team.

Node	Emission	Reception	Status
J5	3117	996	1028.25
J5Gov	1020	1539	639.75
J5Dev	926	1569	623.75
J5Ops	1066	1376	610.5
J52	797	1446	560.75

Reception degree shows that the J5 received less communications than the other members in the team. This is completely normal since he was responsible for initiating most of the communications. All other team members received about the same amount of communications. This is observed because most of the communications that occurred during the planning sessions were destined to all team members rather than being directed to a single individual.

Finally, sociometric status confirms the central role of the J5 within the integrated planning team that was formed for this LOE. Despite the fact that he received less communications than the others, he largely compensated by emitting many more communications than the other members constituting the team. Sociometric status was well-distributed across all other roles within the integrated planning team, suggesting that there was no particular individual dominating the discussions or imposing his/her point of view.

3.6.2 Content Analysis

A content analysis was performed on the communications to assess the relative importance of critical topics during the planning process. The categories of communications were either associated with one of the lines of operation (i.e., security, governance, and development), process-related, team building related, or falling into “others”. A last category was added to take into account inaudible communications.

Number of communications by content type is represented in Figure 42. The first key result is the important focus on process-related communications. These communications are associated to the planning, synchronization, and organization of task work in order to produce the MA brief. They concern the product of the planning team and of the steps required to produce that product. This type of communication was central during the second LOE, probably because the experiment was oriented toward the presentation and the application of a methodology. The large proportion of process-related communications may also be symptomatic of a newly formed integrated planning team. Indeed, since individuals originated from different organisations and had never worked together before, setting up the process and making sure everyone was comfortable with its application should take more time than if the individuals composing the team were all from the same organisation or if they would have worked together in the past. Team building communications, however, did not show up to be very important relatively to other types of communications. This may be caused by the fact that the team members were task oriented (which is desirable) and/or that the team building and handover procedure that

was applied was “sufficient” to build a sufficient sense of teamwork across the members of the integrated team.

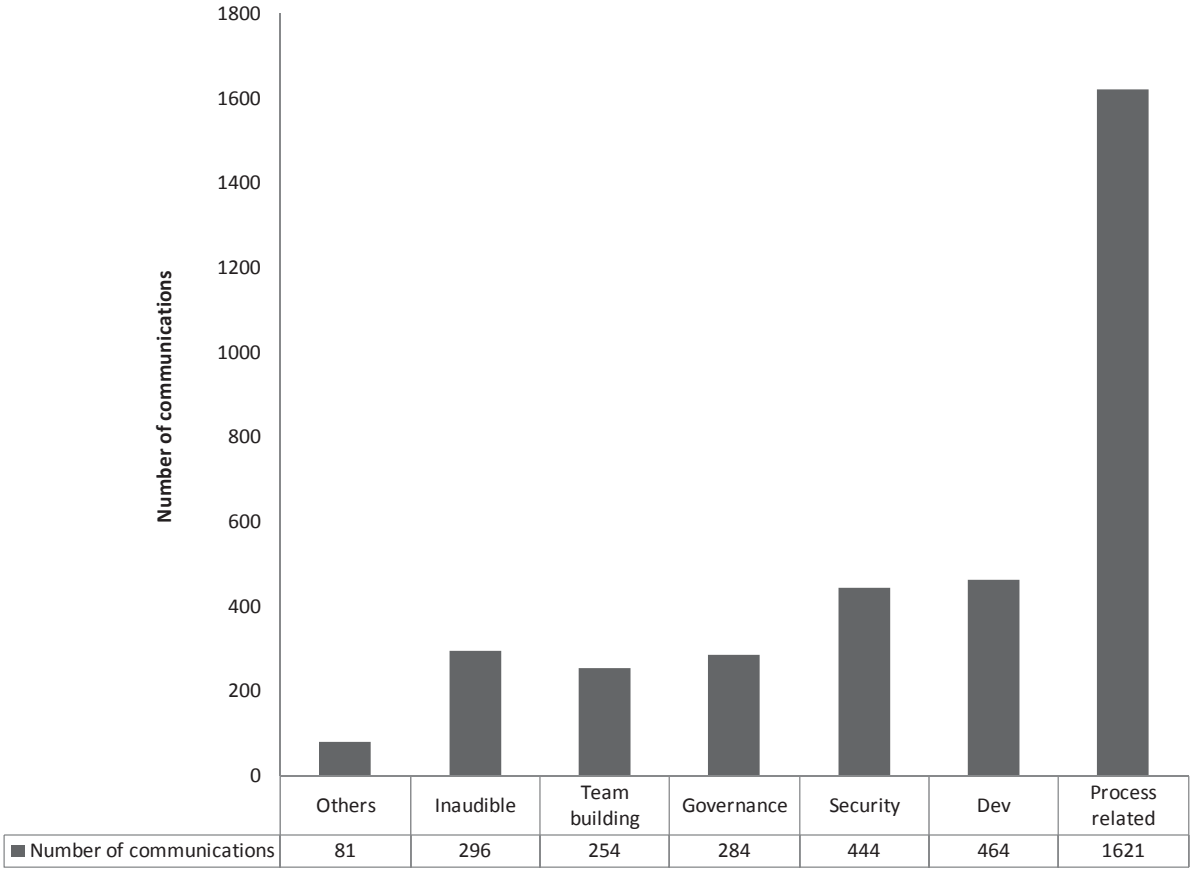
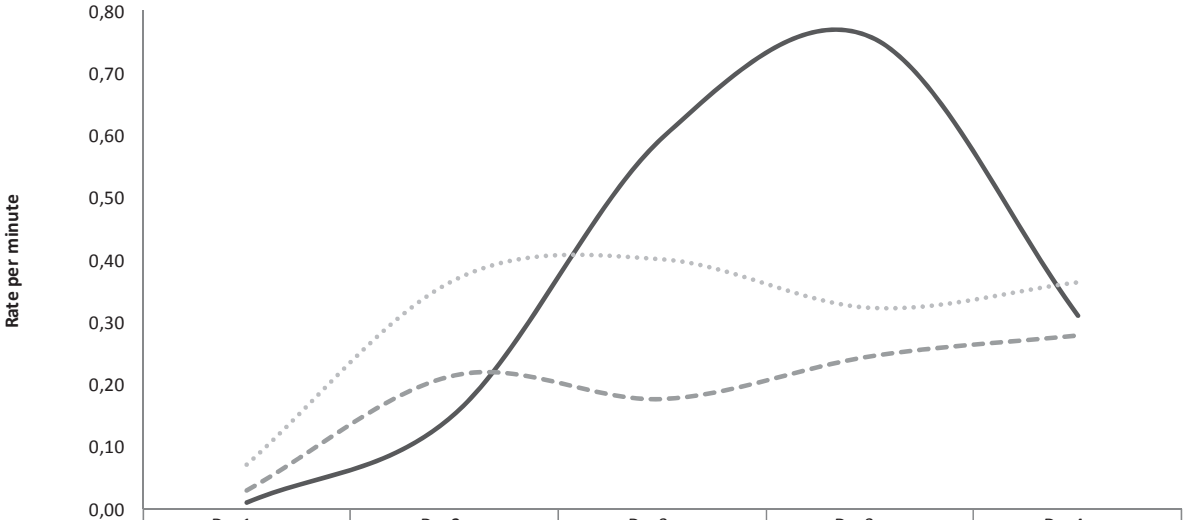


Figure 42. Distribution of communications (entire LOE) across content categories. Process related communications have dominated discussion, followed by development, security, and governance.

Results also show that the number of communications dedicated to security and development lines of operation were equivalent. Communications related to governance, however, were slightly lower compared to the other lines of operation. Figure 43 shows the rate per minute of communications by lines of operation and time period of the LOE. Rate per minute is shown as opposed to absolute frequency because the analysed time periods are not equivalent in length. In other words, by choosing to represent the rate per minute, the time periods are directly comparable even though they did not last the same amount of time. Results show an interesting pattern of communication content through time. The results show that early in the planning cycle, there were less communications and that they were relatively well-balanced across LOO, even though there was a small emphasis on security issues. This was expected since this phase was dedicated for the WoG team to read the scenario material. Results also show that the rate of communications then increases in time concurrently to a shift in content from security to development. The last time period was finally more balanced in terms of communication content. This pattern is interesting

because it shows the shift from a security-driven problem to a development-driven problem coinciding with the announcement of the polio outbreak. The shift was expected because of the shift in the nature of the problem. The observations that development related communications were important during the LOE suggests that the team succeeded (at least during the planning process) to integrate different perspectives in the mission analysis. It demonstrates that the team was capable of shifting priorities as required by the mission.



	Day1 pm	Day2 am	Day3 am	Day3 pm	Day4 am
Dev	0,01	0,15	0,60	0,76	0,31
Gov	0,03	0,22	0,18	0,25	0,28
Security	0,07	0,37	0,40	0,32	0,36

Figure 43. Rate per minute of communications by line of operations and time periods. Rate per minute was used as opposed to absolute frequencies because the length of the time periods analysed was not equivalent from a period to another. By using rate per minute, communications are comparable across time periods. The reported time intervals were all undisturbed mission analysis periods.

3.6.3 Transactive Memory System

In the present study, the TMS questionnaire aimed to assess two subscales of team dynamics: (1) credibility, and (2) coordination. The average ratings of the participants are reported by questionnaire item and sub-scale in Table 21. Results show that both credibility and coordination of the WoG planning team were rated high and very close to the ratings at the same sub-scales during LOE #1. The high credibility score indicates that the SMEs recruited for this study were perceived by their teammates as competent in the OPP. From an experimental point of view, this score reinforces the validity of the results observed. The high coordination score indicates that team members viewed each other as well-coordinated during the exercise. However, we cannot conclude as to whether the high level of credibility and coordination are due to the 12om methodology or to the specific characteristics of the team.

Table 21. Transactive memory systems

Statement	Score (/5)
I was comfortable accepting procedural suggestions from other team members.	4.2
I trusted that other members' knowledge about the project was credible.	4.8
I was confident relying on the information that other team members brought to the discussion.	4.4
When other members gave information, I wanted to double-check it for myself.(reversed)	4.4
In did not have much faith in other members' "expertise". (reversed)	5.0
Credibility	**4.56**
Our team worked together in a well-coordinated fashion.	4.2
Our team had very few misunderstandings about what to do.	3.4
Our team needed to backtrack and start over a lot.(reversed)	4.6
We accomplished the task smoothly and efficiently.	4.0
There was much confusion about how we would accomplish the task.(reversed)	4.0
Coordination	**4.04**
Overall TMS	**4.30**

3.7 Qualitative assessment

Qualitative assessment was performed by component. A general assessment was also performed for any data that would not fit within a specific component. Sources of information included the video and audio recordings of the task-to-tool mapping exercise and of the focus group discussions. They also included the comments reported by the participants and the evaluators on the questionnaires they were handed out. Finally, the notes taken by the observers during the experiment were considered as well.

For each component, content is organized into five categories: (1) Strengths, (2) weaknesses, (3) implementation threats, (4) suggested modifications, and (5) other. This division is based on the component assessment questionnaire, but is also well-suited for the other sources of data. In addition to these main categories, some of them were broken down further to better account for the data. The “Strengths” and “Weaknesses” categories were broken down into “Process” and “Usability”. Process-related strengths and weaknesses refer to information about how a given component will impact on teamwork or the OPP. Usability-related strengths and weaknesses are concerned with specific features or issues with the tool itself. The “implementation threats” category was also broken down into sub-categories. Data was categorized as being “Organisational”, pertaining to the “Maturity level” or “Technological”. Organisational implementation threats are associated with organisational culture. Maturity level concerns tool or process related issues that if not solve would represent a threat to implementation. Finally, technological implementation threats refer mostly to system dependency or the need for IT support. The identification of the sub-categories is data-driven and they reflect the structure that emerged from the analysis of the qualitative data. There are no sub-categories for the suggested modifications and other categories.

Similar comments were not duplicated in the following tables. Moreover, the data shown is in its raw form with very few corrections or clarifications from the analyst. All modifications to the original data are between [brackets]. Finally, French data was translated by the analyst. Translated data is indicated in *italic*. The sources of the data are given to provide context.

3.7.1 WoG OPP handbook

Qualitative data concerning the WoG OPP handbook is summarized in Table 22. Qualitative assessment suggests that the OPP handbook should not constitute a rigid guideline, but rather a general approach to planning. Moreover, participants did not agree between each other whether the handbook should integrate more civilian planning processes or stay focused on the military OPP. Overall, however, participants all agreed that the material was good enough in its actual shape to be useful in operational context.

Table 22. Summary of qualitative data on WoG OPP handbook

Strengths	
Comment	Source

• Process
  ○ Good reference material.
  ○ More for a non-initiated (to OPP) planner.
  ○ A good description of process + neutral vocabulary.
• Usability
  ○ Common denominator, low on acronyms, minimal staff process branching out of main cycle.
	Component questionnaire (Appendix B)

Weaknesses	
Comment	Source

• Process
  ○ Too rigid in application (details doctrine).
  ○ Military/linear biases.
	Component questionnaire (Appendix B)

• Process ○ Both the glossary and the handbook look like the CFOPP made comprehensible for civilian rather than an integration of the civilian planning style. [02:46:00]	Focus group discussion [time in recording]
Implementation threats	
Comment	Source

• Organisational
  ○ Not in the civilian /DFATD way of doing/seeing things more difficult to apply on our side.
• Maturity level
  ○ Work needs to bridge OPP with management by systems in civilian planning (not necessarily RBM) typically from TBS or academic project management too.
• Technological
	Component questionnaire (Appendix B)

Suggested modifications	
Comment	Source
• Highlight differences that a well-versed-in-OPP planner should be aware of. • Given a real situation/training it gives a common ground that should be discuss by lead planner of each GoC Dept to fully reach effort and expectation before actual planning start. • An experienced team composed of DFATD and CAF should develop a process that meets the needs of both synchronizations, dealing with the meaning of OPP + RBM where possible.	Component questionnaire (Appendix B)
• [Integrate civilian planning methods to the document] [02:46:30] ○ [however] We are in a military exercise and civilians are integrated within the exercise [02:50:40] • Highlight what is different from the CFOPP [03:02:30] • People should read the handbook and then the J5 should present the OPP briefly to the integrated team. [03:08:59]	Focus group discussion [time in recording]
Other comments	
Comment	Source
• Short and to the point. Could be bias by the fact that I already knew the process. • A handbook is a useful tool, but should be the basis for group discussion and practical exercise to ensure understanding.	Component questionnaire (Appendix B)
• *OPP handbook would be useful for everything [tasks of the OPP] but has a low utility [01:00:25]*	Task-to-tool mapping discussions [time in recording]

3.7.2 Team building and handover procedure

Qualitative data concerning the team building and handover procedure is summarized in Table 23. Qualitative analyses show that participants were generally satisfied with the team building and handover procedure. Implementation threat included time required to perform the exercise and willingness to participate.

Table 23. Summary of qualitative data on team building and handover procedure

Strengths	
Comment	Source

• Process
  ○ Underlines different perspectives.
  ○ Focus the common understanding of the problem.
  ○ Can take very little time.
  ○ Ice breaker, permits individual staff to share strengths w/o seeming defensive/arrogant.
  ○ [team building and handover procedure] Forces people to tell others what they think is their own task.
	Component questionnaire (Appendix B)

○ Allows the team to think in an integrated manner in the early stages of planning, clarify roles + responsibilities + reporting structures ○ Formalizing an informal process that may or may not occur. ○ Especially useful with a new team or a turbulent staffing environment. ○ Helps address the problem of absences. • Usability	
Weaknesses	
Comment	Source
• Process ○ Can influence others. ○ Somewhat dependant on personalities and seriousness accorded by participants. ○ For a JIMP environment it may be very DND acronyms or concept heavy. ○ Planning is not always linear in practice; officers are quick to discuss practicalities/operations. • Usability ○ Excel based vs. more easy to read or visualize. ○ Table format not ideal. ○ Roundtable with scribe rather than filling spreadsheet.	Component questionnaire (Appendix B)
Implementation threats	
Comment	Source
• Organisational ○ Time required. ○ Presence of all involved members. ○ Willingness/open-mindedness (cultural resistance to change). ○ Every officer is ultimately accountable through different lines/dept (ex. DFATD vs. DND). • Maturity level • Technological	Component questionnaire (Appendix B)
Suggested modifications	
Comment	Source
• Include a more personal aspect to permit linking. • Discipline of participants. • Chart of roles and responsibilities could be outlined in writing. • Rather than expected departure date, it might be better to use anticipated absences to cover for other duties, leave, etc. • Add reporting chain; include tasks to each planner (contribution).	Component questionnaire (Appendix B)
Other comments	
Comment	Source
• Part I is covered via other means normally. Different road, same results. Part II is really a nice upgrade.	Component questionnaire (Appendix B)

3.7.3 Interactive common glossary

Qualitative data concerning the interactive common glossary is summarized in Table 24. Participants were generally favourable toward the common interactive glossary as it provided them a common place where to share definitions. They found it was easy to use and good reference material. There were some issues mentioned such as (a) the possibility to duplicate terms, (b) the definitions could change in time (or depending on the commander) and (c) it was lacking civilian terms as it was presented in the study. On the other hand, they mentioned it was easy to amend. Overall, in order to solve these issues, participants agreed that cooperation between departments to build a more comprehensive glossary is critical.

Table 24. Summary of qualitative data on interactive common glossary.

Strengths	
Comment	Source

• Process
  ○ Writes a common place all definitions.
  ○ Provides a shared interpretation of each term.
  ○ Definitions should be explanatory, where official versions sometimes are not.
  ○ Good starting point to compare and present meaning of terminology.
• Usability
  ○ Good reference material.
  ○ Easy to use.
  ○ Glossary can be easily amended.
	Component questionnaire (Appendix B)

• Process
  ○ Someone who is not familiar with OPP will appreciate the glossary [02:45:40]
  ○ Common ground to start and then expand [02:55:00]
	Focus group discussion [time in recording]

Weaknesses	
Comment	Source

• Process
  ○ Can sometimes change the meaning of things.
  ○ Can lead to repetition of terms.
  ○ Civilians without formal training in project management may not be familiar enough with the significance of certain new concepts to fully grasp meaning and differences between OPP language and civilian equivalents.
• Usability
  ○ Too long glossary can be confusing.
	Component questionnaire (Appendix B)

• Process ○ Both the glossary and the handbook look like the CFOPP made	Focus group discussion [time in recording]

comprehensible for civilian rather than an integration of the civilian planning style. [02:46:00]	
Implementation threats	
Comment	Source
• Organisational ○ Risk of cultural push track if lead for future drafting not shifted to DFATD. • Maturity level • Technological	Component questionnaire (Appendix B)
Suggested modifications	
Comment	Source
• Add criteria of success. • Should identify the source of the definition. • Consultation need to continue between departments to build more visibility and acceptance.	Component questionnaire (Appendix B)
• [Integrate civilian planning methods to the glossary] [02:46:30]	Focus group discussion [time in recording]
Other comments	
Comment	Source
• This handbook and glossary need to be part of a broader inter-agency arrangement for training future complex stability deployment.	Component questionnaire (Appendix B)
• *We already have a dictionary, but it doesn't necessarily integrate governance issues. Therefore, the interactive common glossary provides an integrated component. [01:02:15]*	Task-to-tool mapping discussions [time in recording]
• Potential issue with changing definitions through time (i.e., when commander changes) [02:56:30] • Potential issue with multiple definitions of the same term [02:57:30]	Focus group discussion [time in recording]

3.7.4 Collaborative knowledge representation

Qualitative data concerning the collaborative knowledge representation is summarized in Table 25. Participants were satisfied with the knowledge representation tool. Its main advantages included (a) the possibility for each member to explicit his/her knowledge of the situation (especially in individual knowledge representation), (b) the maximization of knowledge shared, and more generally (c) the thought process it generates during the completion of a graph. Participants generally found the software fast, reliable, and intuitive. They also mentioned that this tool could easily be exported outside OPP. The main concern is that the creation of conceptual diagrams is time consuming. One interesting suggestion for improvement is to establish a set of guidelines for the creation of conceptual diagrams. This could be done to make sure the individual graphs are more similar and consequently easier to merge afterward. For instance, what is the level of the analysis? How many concepts should we focus on?

Table 25. Summary of qualitative data on collaborative knowledge representation.

Strengths	
Comment	Source
• Process ○ Increases the inputs of each participant. ○ Produces a usable product that will be inserted into MA brief. ○ Following Ma brief, allowed us to come-up with subsequent slides (i.e. essential tasks, risks, implied task and OP Design). ○ Maximizes sharing of ideas + knowledge. ○ Refine though process; make you think about more concepts/influences that you start with. ○ Very useful to deconstruct our thoughts and point out the most important concepts and their real impact. ○ If used in early stages, can help the team to clear out unimportant concepts and focus on important points from various perspectives. ○ If intention is to move on to single diagram representing inputs of all individuals, then best to give/agree on parameters (e.g. scale, level of details) and perhaps major concepts, issues, threats, etc. ○ Conceptual diagrams can help bridge the RBM/civilian method of problem tree, or strength, weaknesses, opportunities, threats development civilians use + support common understanding. • Usability ○ Fast, stable, flexible, involving. ○ Fast learning, intuitive, stable, easy to standardize same concept-relation. ○ Limited usability as long as it is not shared and compared with the other members of the team.	Component questionnaire (Appendix B)
• Process ○ *Cross-impact method and collaborative knowledge [representation] are the two most exportable components [outside of OPP]. [01:06:20]*	Task-to-tool mapping discussions [time in recording]
Weaknesses	
Comment	Source
• Process ○ Less vocal/confident members' views may not be sufficiently voiced/shared. ○ Doing this individually is not profitable - the benefit of this analysis is the exchange of information and gaining a common	Component questionnaire (Appendix B)

understanding. ○ Names of concepts and types of relations are without limits. • Usability ○ Time consuming ○ Complicated and requires training.	
Implementation threats	
Comment	Source
• Organisational ○ Initial resistance to the investment to use new tool. ○ Leadership cost to maintain use of tool and continued use in planning cycle. • Maturity level ○ Scope of analysis - too many things can be put in a common diagram. ○ Computer issues (loss of data, etc.). • Technological ○ Resources (time, computers, power, IT support).	Component questionnaire (Appendix B)
Suggested modifications	
Comment	Source
• Export of chart in a stable format by doing query of content and relations. • Prepare a little guide to use the software. • Establish a limit at some point on the list of concepts and relations in order to facilitate integration and common use of the tool.	Component questionnaire (Appendix B)
• *It would be nice to see the differences between graphs. [01:13:24]* • *Software could find opposing ideas (or conflicts) between graphs [01:14:02]*	Task-to-tool mapping discussions [time in recording]
Other comments	
Comment	Source
• Could foster cohesion and buy-in within a JIMP environment as this is neutral ground i.e. not DND doctrine. • The cross-function conceptual diagram is very useful as a joint process - not as an individual task.	Component questionnaire (Appendix B)
• In my opinion the diagram produced collaboratively did not demonstrate the common understanding.	Process evaluation (Appendix I)
• *[Collaborative knowledge representation] is the tool we used the most and the one that generated the most discussions.[00:58:50]* • *[Collaborative knowledge representation] generates discussions whereas [cross-impact method] is more "practical".[00:59:25]* • *Collaborative knowledge representation is what made us think about the situation [00:59:50]* • *If I had to choose three components for planning, it would be collaborative knowledge representation, [OP Design tool], and the*	Task-to-tool mapping discussions [time in recording]

cross-impact method. [01:01:36] • *Cross-impact method and collaborative knowledge [representation] are the two most exportable components [outside of OPP]. [01:06:20]*	
• I believe that KR and CIM could help tremendously campaign tracking [03:33:10]	Focus group discussion [time in recording]

3.7.5 Cross-impact method

Qualitative data concerning the cross-impact method is summarized in Table 26. CIM was deemed very useful by most of the participants, although one civilian put a caveat on its utility from his/her perspective. Its benefits are mostly attributed to the rigor and systematic approach it enforces in the analysis of key factors. The output that it generates was also especially appreciated, notably the assessment of second and third order effects. The main drawback of the CIM is the complexity of the analysis (it's hard to understand) and the time it takes to complete. It is still unclear based on the comments of the participants whether the CIM should be used in stage 2 or stage 3 of the OPP. In order to improve the method, it would be helpful to better define what a factor is in this context.

Table 26. Summary of qualitative data on cross-impact method.

Strengths	
Comment	Source

• Process
  ○ Very useful.
  ○ Gives great results and representation of our thoughts.
  ○ Easy to share.
  ○ [In the context of] longer planning cycles, larger and less familiar planning teams, less familiar problems. Could reveal important knowledge that was not yet been shared.
  ○ Allows for a more formal methodology to rank factors and uncover linkage between factors to confirm identification of key problems and underlying contributing factors
• Usability
  ○ Easy to use.
	Component questionnaire (Appendix B)

• Process ○ How effectively did the team identify objectives? → in discussion around cross-impact analysis	Process evaluation (Appendix I)
• Process ○ *Cross-impact method and collaborative knowledge [representation] are the two most exportable components [outside of OPP]. [01:06:20]*	Task-to-tool mapping discussions [time in recording]
• Process ○ This [CIM] brings a lot of honesty toward if your COAs feasible (sig) like is my COA robust enough to go. This [CIM] brings more rigors to your COA [02:37:30]	Focus group discussion [time in recording]

○ There is a lot of thought power that goes in this […] when you face a complex situation […] mapping knowledge then CIM before COAs is useful [02:39:30] ○ CIM is a good thing (sig) because it gives you 2nd and 3rd order effects which are things that are hard to capture through normal OPP if you don't use [CIM] [02:41:50] ○ The final "matrix view" is very "cool" [02:43:25]	
Weaknesses	
Comment	Source
• Process ○ Difficult to understand. ○ Subjective weighting could undermine the validity of what appears to be objective results. ○ Quantifies subjective assessment of importance of factors and their relations that are still open to subjective biases and errors. • Usability ○ Time consuming if done thoroughly. ○ Requires substantial investment in training.	Component questionnaire (Appendix B)
• Process ○ Issue with the neutral answers which might have had an impact on the end result. [02:43:10] ▪ The factors are so interrelated that it is not as easy to say that "Factor x" will lead to an increase in "Factor Y". [The reality is that] it's not that much of a direct relation. [02:44:10]	Focus group discussion [time in recording]
Implementation threats	
Comment	Source
• Organisational ○ Can be difficult to reach consensus on individual factors ○ Because of level of effort in training and time required during OPP process to complete collaboratively, risk that tool would be skipped in the absence of leadership to use it. • Technological ○ System dependant	Component questionnaire (Appendix B)
Suggested modifications	
Comment	Source
• Factors are concrete things that you can touch, they're not concepts [02:40:00] • I would put the CIM in stage 2 no into stage 3 [02:40:30] ○ But you need to have done your graphs before [02:42:10] • You need to come up with a functional definition of the concepts [02:41:40]	Focus group discussion [time in recording]
Other comments	
Comment	Source

• Could be useful in strategy/institutional planning. • Rated this highly because it provided a very useful result -> the primacy of security. • Tool probably ideal for campaign design and analysing complex situation than for comprehensive OPP cycles for branch plans on limited fragOP in time and space.	Component questionnaire (Appendix B)
• *[Collaborative knowledge representation] generates discussions whereas [cross-impact method] is more "practical".[00:59:25]* • *[Cross-impact method] allows you to see the sequence of actions to perform [00:59:38]* • *We were able to weight the importance of the factors with the cross-impact method [01:00:10]* • *If I had to choose three components for planning, it would be collaborative knowledge representation, [OP Design tool], and the cross-impact method. [01:01:36]* • *Cross-impact method and collaborative knowledge [representation] are the two most exportable components [outside of OPP]. [01:06:20]*	Task-to-tool mapping discussions [time in recording]
• I believe that KR and CIM could help tremendously campaign tracking [03:33:10]	Focus group discussion [time in recording]

3.7.6 OP Design tool

Qualitative data concerning the OP Design tool is summarized in Table 27. The OP Design tool was also very much appreciated, especially by the military team members. Its main strength is that it is tailored for OPP and consequently very useful in this context. However, its main strength is also its main weakness, as the civilians felt that it did not fully capture their way of understanding the situation. Most of the improvements that could be done that were specified are minor and technical. The process itself was satisfying to most/all integrated team members.

Table 27. Summary of qualitative data on OP Design tool.

Strengths	
Comment	Source

• Process
  ○ Cannot be more task-tailored.
  ○ Understanding and definition of the problem.
  ○ Good representation.
• Usability
  ○ Good representation; easy to comprehend & use.
  ○ Standardization of visualization of OP design.
  ○ Possibility to copy/paste from other software.
	Component questionnaire (Appendix B)

• Process ○ How effectively did the team consider the impact of time on	Process evaluation (Appendix I)

the operation? → The operation Design tool timeline demonstrated a comprehensive consideration of this element ○ How effectively did the team identify objectives? → Displayed on Op Design tool ○ How effectively did the team describe the desired end state? → Displayed on Op Design tool	
• Process ○ Very useful to outline key task/activities that need to happen from different LOO [02:10:45] ○ Synchronize assets that only INT can have [02:12:12] ○ Fast/stable, all required box are there [02:13:00] ○ You can work on it with other product [02:15:00] ○ You keep track of each changes as you make them [02:15:15] ○ Undeniably essential to the OPP [02:15:45] ○ OP Design forces you to be sequential [02:26:10]	Focus group discussion [time in recording]
Weaknesses	
Comment	**Source**
• Process ○ The use of timelines can encourage linear, sequential thinking, when what may be required is more concurrent activities. ○ Could make more distinction between continuous and punctual priorities/actions. ○ Does not capture unexpected events/disruptions in the planned cause of action. ○ Not in our institutional habits to plan on such a short period of time (civilians). Very difficult activity for civilians from DFATD ○ May result in detailed synchronization of tasks (vs DPs) • Usability ○ Requires system & software available ○ Software development required	Component questionnaire (Appendix B)
• Process ○ *DP analysis is tightly linked to OPP, I would not use it in another context. [01:06:30]*	Task-to-tool mapping discussions [time in recording]
• Process ○ Not as easy to capture ongoing considerations throughout. It is more oriented toward tasks than results [02:10:55] ○ Three LOOs is a constraint that we should not follow. We could have up to six LOOs for instance. [02:20:20] • Usability ○ Only weakness are technicalities [02:13:30] ○ Technical issues must be solved to be used in operational context [02:16:45]	Focus group discussion [time in recording]
Implementation threats	
Comment	**Source**
• Organisational ○ There will be resistance to adding any tool that is not deemed	Component questionnaire

essential. ○ Org. Cultural way of doing thing and time window can lead to conflictual discussion and planning effort. • Maturity level ○ Flexibility of visualization (little manipulation impossible) • Technological ○ Only tool takes time to implement, memory on computers, etc.	(Appendix B)
Suggested modifications	
Comment	**Source**
• Decision point should be represented by yellow star. • Needs to be able to have simultaneous DPs. • Connecting lines between LOOs to one DP should be possible. • Objectives should be able to be joined to multiple LOO. • Allow preconditions to be identified in a box to left of Op Design • When "return" is used, text should fall on next lines. • Arrows to indicate DP is ongoing. • Add importing feature from CIM.	Component questionnaire (Appendix B)
• Should be able to number the DPs [02:21:00] • Should be able to have on each DP the measure of effectiveness, the criteria for success and the indicators that are related to that [02:21:20] • Would you see the tool as a possible monitoring tool? – Definitely. [02:22:02]	Focus group discussion [time in recording]
Other comments	
Comment	**Source**
• I think that in real world, given more time and better understanding of each org culture it will be a successful process. • Should be a concerted effort to move away from security dev and gov as the fixed LOO as likely counterproductive in analysing complex problems instead of focusing on institutional view.	Component questionnaire (Appendix B)
• *OP Design is a good representation of the problem, but the OP design doesn't make you think about the problem. [01:01:20]* • *[mil to civ] I would have thought that the tool was generic enough that you would like to use it in your context [01:06:50]* ○ *[civ to mil] It lacks an overview of the continuity of interventions [01:07:00]* ○ *[civ to mil] it's to constrained in time [01:07:10]* • *If I had to choose three components for planning, it would be collaborative knowledge representation, [OP Design tool], and the cross-impact method. [01:01:36]* • *[OP Design] is tightly linked to OPP, I would not use it in another context. [01:06:30]*	Task-to-tool mapping discussions [time in recording]
• This should be deployed and used as we speak as far as I'm concerned [02:22:55]	Focus group discussion [time in

• OP Design should happen after CIM [02:34:30]	recording]

3.7.7 Integrated MA briefing template

Qualitative data concerning the integrated MA briefing template is summarized in Table 28. The integrated MA briefing template was appreciated, but mostly considered as reference material rather than a tool. Participants underlined the importance of specifying that the template is a general guideline rather than a pre-determined format that you have to follow. The main potential issue associated with it is to omit an important factor because its category was missing from the template. Consequently, it is critical that the template is comprehensive and extensive.

Table 28. Summary of qualitative data on integrated MA briefing template.

Strengths	
Comment	Source

• Process
  ○ Forces integration of all domains (gov/dev/sec).
  ▪ Reinforces the requirement to report different perspectives to the leadership.
  ○ Synch of supporting activities.
  ○ Provides a standing example of what could constitute a comprehensive WoG briefing.
• Usability
  ○ Easy to use/fill.
  ○ Visually efficient.
  ○ Comprehensive, logic.
	Component questionnaire (Appendix B)

• I would use it as a point of departure. [03:11:00]	Focus group discussion [time in recording]
Weaknesses	
Comment	Source

• Process
  ○ Consistent division of ppt law dev/gov/CAF will reinforce distinction/differences. An indication of LOO would be more appropriate (dev/gov/sec, potentially).
  ○ Entire format may not be relevant to the problem; time spent in filling out slides may/will take time away from essential points of the briefing.
  ○ Planning teams could use the template as a guide to the conduct of the OPP, which could lead to the omission of important steps.
  ▪ Some factors were omitted (iPOE, terrain, time and space, troops to task).
• Usability
  ○ Too long
	Component questionnaire (Appendix B)

• Process ○ The use of the format for briefing tended to drive the contents – they did not adapt the format to the requirement, probably because we told them to use it.	Brief evaluation (Appendix J)

• Process
  ○ May reinforce the separation of the problem into LOO [03:12:00]
  ○ You assume that the lines of operations are defined at this point [03:13:00]
	Focus group discussion [time in recording]

Implementation threats	
Comment	Source

• Organisational
  ○ Resistance to change should be minimal (some may prefer to use already existing tailored ppt).
  ○ Different HQ will tend to use their own formats, but it is still worth the effort.
• Maturity level
  ○ Template would have to be updated with OPP manual.
  ○ Needs detailed analysis to support input to the briefing.
• Technological
  ○ System dependant.
	Component questionnaire (Appendix B)

Suggested modifications	
Comment	Source
• Move conceptual diagram to factors analysis. • OP Design template added. • Change governance for political. • The template should be incorporated into the OPP manual; more likely to be used. • Add missing factors. • Work collaboratively to build the briefing. • Shorten format to key factors, deduction, risks and OP design - focus on core elements of the problem as jointly perceived by the team and what comd/roc would have as potential adhesions.	Component questionnaire (Appendix B)
• *It's nice if the template allows for flexibility. [01:05:34]*	Task-to-tool mapping discussions [time in recording]
• Make sure the list of factor included in the template is extensive. [03:15:00] • It needs to be adapted by the commander for the specific situation [03:17:50]	Focus group discussion [time in recording]
Other comments	
Comment	Source
• A great point of departure… format will evolve based on comd preferences and JIOPG preferred practices.	Component questionnaire

• Very similar to the one already in use in HQ.	(Appendix B)
• *The template represents the sequence of OPP. So it gives a common ground on what is needed. It allows to do OPP without having to comprehensively understand OPP [01:04:39]* • *Can be used as a basis for any planning process, not only OPP. [01:05:20]*	Task-to-tool mapping discussions [time in recording]
• [The process was not template driven, it was the thought process done before that allowed the team to fill the template rapidly] [03:14:00]	Focus group discussion [time in recording]

3.7.8 General

Some of the qualitative data did not refer to any particular component. Two categories were identified. The first category refers to the tools in general (or 12om methodology) rather than being specific to a component. The second category refers to the experimental setup itself, mostly in terms of the scenario material.

Tools

We were missing a tool to conduct "troop to task" planning [01:18:00].
All tools presented to us serve a purpose [03:28:00]
The tools fit at a high strategic JIMP level [03:29:35]
They don't fit at brigade level [03:29:40]

Experimental setup

Experimental related content
The read-in needs to happen before [03:21:00]
The [scenario] content was more than sufficient [03:23:00]

Overall, these comments can be informative either for the deployment of the 12om methodology, for future work or future experimentations.

4 Conclusions and Recommendations

The 12om project seeks to develop a *methodology* to improve the understanding of a complex situation by a multidisciplinary, government-wide team. The specific objectives of the current study were to assess the 12om methodology (and its components) in terms of its potential for supporting the collaboration processes, integration of different perspectives, and common understanding. Complementary measures to help capture the outcomes of LOE #2 were linked to performance, process and product evaluation.

4.1 Key Observations and Associated Recommendations

Overall, the 12om methodology was rated very favourably. Results of the E-MYRIAD analysis indicate that the 12om methodology, as used in LOE #2, was 80% successful. It must be noted however, that the E-MYRIAD analysis assesses the proximity of an ideal outcome as defined by the analyst team and may not perfectly reflect the opinion of SMEs or stakeholders. Another E-MYRIAD analysis could be performed with the preference model of a different stakeholder and generate qualitatively and quantitatively different results. That being said, the actual analysis reveals that the main strength of the 12om methodology is in supporting the three dimensions of Collaborative Understanding in a very balanced way. Results in terms of feasibility are relatively good but somewhat lower, mainly due to training time/effort which is seen as demanding and partly due to use time/effort which is seen as somewhat demanding as well.

In terms of support to the three dimensions of Collaborative Understanding (i.e., Collaboration process, Integration of different perspectives, and Development of a common understanding), the 12om methodology has performed very well. Even though ratings were very positive, E-MYRIAD shows that the greatest value in improving a single metric would come from a change that would impact the Objective *Development of a Common Understanding* metric. In other words, the development of the 12om methodology should prioritize the improvement of this particular dimension.

Although 12om would benefit the most from improving the development of a common understanding, the results show that this particular dimension was objectively very high. Indeed, correlations of the "mental models" between participants were very high (especially for participants who used the CIM component). In other words, the support to the development of a common understanding from 12om is very high, and would benefit from further improvement.

Another interesting finding is that the toolbox of components covers a wide area of support in terms of OPP tasks and dimension of collaborative understanding. Indeed, functional gap analysis reveals that there was no task or dimension of collaborative understanding left unsupported, mainly because of the flexibility of some of the component (notably *WoG OPP handbook* and *interactive common glossary*). Moreover, the analysis reveals that the high impact on task work of some of the components confer them a great value within the 12om toolbox.

4.1.1 Performance, Process, and Product

Evaluations of the MA brief made by the SME observers were relatively good, although there is room for improvement with that matter. Overall, evaluations indicate that the product of the OPP generated by the WoG team met SME's expectations, but did not exceed them. This may be due to the nature of the LOE, which was mostly focused on training and using the 12om components, rather than conducting a thorough mission analysis. Other measures (not directly associated with the product but related to the process) show that the team was working in an "optimal zone" in terms of workload and SA. Indeed, levels of workload and SA, respectively measured by the NASA-TLX and MARS, are good indications that the team was efficient.

Moreover, the team appeared to have completed the task seriously and with a high level of engagement. Indeed, the team dynamics analyses all point toward great team cohesion, collaboration, and seriousness. Social network analysis clearly shows the active leading role of the J5 (which was expected). Moreover, the analysis suggests a high level of collaboration between participants (and most importantly across lines of operations). All participants save J5 had similar sociometric status, emission degrees and reception degrees, revealing similar levels of participation during the exercise. Content analysis suggests that members of the WoG team made efforts for integrating different perspectives during the process since the development-related content (which is often left apart in "integrated" operations) was by far the most important topic at some point during the LOE. Finally, TMS shows that the members of the WoG team perceived the other team members as being credible and found that coordination within the team was good.

4.1.2 Individual components

Overall, all components were rated to be useful. Only two components were rated below 6/10 by the SME observers (i.e., *OP design tool* and *creation of the common vocabulary*), and none was rated below 6/10 by the WoG team. General time and effort required to use the components was low, even though there is room for improvement with that matter. Finally, most components and sub-components were rated as having benefits that outweighed the costs. The lowest ratings were associated with IMAGEv3 tool sub-components: *sharing conceptual diagrams* and *querying conceptual diagrams using filters*. Taken together, these results suggest that the core features of the 12om methodology are good enough to be used by most users, but that the more complex, specialized features should be trained further or delegated to "technicians".

Having in mind the limitations of the E-MYRIAD analysis, the combining of disparate measurements into a meaningful aggregate assessment of the 12om methodology and its components showed that the main outcome of LOE #2 is that the current toolset was deemed very satisfactory (80%), with a support of collaborative understanding of 86% and an overall feasibility of 73%. The main challenge identified by this analysis was the overall need to address the issue of training time/effort. Fortunately, this is an issue that can be readily addressed by optimizing training time/effort using a common example for all components and by developing an interactive multimedia tutorial to efficiently show how to use each component.

Individual component assessments varied between 63% and 82% showing that there is still room for improvement. A sensitivity analysis based on the MYRIAD models helped identify the most valuable areas for improving each component. Results for the twelve components and sub-components are summarized in Table 29, which show not only the MYRIAD score but also the cost-benefit ratio ratings of SMEs and, most importantly, the average rating on only the three fundamental objectives of the 12om project, that combined are referred to as "collaborative understanding support").

Table 29. MYRIAD output for the 12om methodology components analysis

Component	MYRIAD Score	Cost-Benefit Ratio /10	Collaborative Understanding Support /10
OP Design Tool	82%	7.7	7.5
Conceptual Diagrams (Collaborative)	76%	8.7	8.8
Team Building and Handover Procedure	76%	8.4	8.1
Common Glossary	76%	8.0	8.0
OPP Handbook	71%	7.4	7.3
WoG MA Brief Template	70%	8.1	7.0
Conceptual Diagrams (Individual)	68%	6.4	6.7
CIM	63%	7.5	7.3
Sub-component	**Score**		
Creating a Common Vocabulary	81%	7.8	7.9
Sharing Conceptual Diagrams	69%	5.0	7.2
Creating Views Using Filters	68%	6.0	6.4
Querying Using Filters	68%	5.4	7.0

The Collaborative Understanding Support Metric is focused only on the priority objectives of the 12om project, and therefore not comparable to the other two metrics. In comparison to MYRIAD assessments, a somewhat different portrait was obtained by directly asking participants to rate the cost-benefit ratio of each component. Based solely on this measure the Collaborative Conceptual Diagrams component stood out as being the most valuable. Nonetheless, the average cost-benefit rating for the main components of the 12om methodology was 7.78/10, which closely matches the overall MYRIAD assessment of 80%. The correlation between the cost-benefit rating of participants for the 12 components and sub-components and the MYRIAD assessments is statistically significant, but only of moderate strength, $r(10)=.567$, $p=.027$ (1-tailed). This result illustrates that the preference model better reflects the analysts' perspective rather than the SMEs' perspective. However,

it should be noted that the reliability of a single questionnaire item as an assessment of a component is necessary limited as well. Also, it should be noted that is not the individual value of each component that should be the primary concern but rather their combined effectiveness as a complementary toolset.

That being said, it is still possible to improve individual components of the 12om methodology in order to increase its overall support or feasibility. Specific SME recommendations per component are summarized in Section 3.7 and will not be repeated here. The MYRIAD sensitivity analysis allowed identifying which areas to prioritize. Below we summarize the key findings from the MYRIAD multi-criteria assessment and the flexibility/specialization analysis, and identify what would be the most valuable areas to improve according to the MYRIAD sensitivity analysis.

4.1.2.1 OPP Handbook

Results indicate that the OPP Handbook is 71%atisfactory according to the multi-criteria analysis. Results in terms of feasibility are relatively good. Results also indicate that the OPP handbook is a very flexible component, useful throughout the whole OPP process.

According to the sensitivity analysis, the key areas for improvement are any of the five dimensions of support (deemed equally beneficial):

- Impact on taskwork;
- OPP Support;
- Integration of Perspectives;
- Development of Common Understanding; and
- Support to Collaboration.

4.1.2.2 Team Building

Results show that Team Building component is 76% satisfactory according to the multi-criteria analysis. Its main weakness is its relatively low impact on taskwork, since it is mainly categorised as reference material. Otherwise, with the exception of a moderately demanding use time/effort, results are extremely favorable. Interestingly, this component does not support directly any sub-tasks of the OPP, as its main benefits are associated with the support of the three dimensions that 12om aims to support (i.e., Integration of different perspectives, collaboration, and shared understanding). This suggest that this component could easily be adapted to fit other contexts (apart OPP) involving integrated teams.

According to the sensitivity analysis, the priority areas for improvement are:

- OPP support;
- Integration of perspectives;
- Common understanding; and
- Impact on taskwork.

4.1.2.3 Common Glossary

Results indicate that the Common Glossary is 76% satisfactory according to the multicriteria analysis. Its sole weakness is its relatively low impact on taskwork, since it is mainly categorised as reference material. This component was one of the most flexible from the 12om toolbox as the participants considered it to support all sub-tasks of the OPP from initiation to COA development.

According to the sensitivity analysis, the most valuable areas for improvement are:

- Impact on taskwork;
- OPP support;
- Common understanding; and
- Collaboration support.

4.1.2.4 Conceptual Diagrams (Individual)

Results from the multi-criteria analysis indicate that this component is 68% satisfactory according to the multi-criteria analysis. Its main weakness is related to feasibility in terms of time and effort required for training and use. However, despite these difficulties, this component is deemed highly valuable especially because of its important impact on the taskwork to perform. Moreover, this component is relatively flexible, providing support to a wide range of sub-tasks of the OPP.

According to the sensitivity analysis, the key areas for improvement are:

- Improve any one of dimension of support (other than impact on taskwork which cannot be further improved);
- Reduce use time/effort (e.g., place a time limit on list of concepts/relations); and
- Reduce training time/effort.

4.1.2.5 Conceptual Diagrams (Collaborative)

Results from the multi-criteria analysis show that this component is 76% satisfactory according to the multi-criteria analysis. Its main weakness relates to feasibility, specifically its demanding use time/effort. Nonetheless, this component is deemed highly valuable because of its high positive impact on the taskwork, and is the component with the highest (94%) satisfaction in terms of *Dimensions of Support*. Moreover, this component is relatively flexible, providing support to a wide range of sub-tasks of the OPP.

According to the sensitivity analysis, the key areas for improvement are:

- Reduce use time/effort; and
  Reduce training time/effort.

4.1.2.6 Cross-impact method

As assessed in LOE #2, CIM component is 63% satisfactory according to the multi-criteria analysis. Its main weakness is feasibility, specifically in terms of its use time/effort.

Nonetheless, this component is deemed highly valuable in terms of its impact on the taskwork and received quite high usability ratings when considering the nature of the analytical method.

According to the sensitivity analysis several possible areas for improving the cross-impact method are equally viable:

- Reduce use time/effort;
- Increase OPP Support;
- Improve Integration of Perspectives;
- Increase support to Development of Common Understanding;
- Improve support to Collaboration; and
- Reduce training time/effort.

The choice may therefore go to an area that can be more easily improved by the design team, such as optimising training time/effort using a multimedia tutorial or redesigning part of the process to reduce use time/effort.

Moreover, with regard to the moment of use of this component, the functional gap analysis suggests that it should be used during mission analysis rather than during initial COA development.

4.1.2.7 OP Design Tool

OP Design Tool component is 82% satisfactory based on the MYRIAD multi-criteria analysis. *It is the most successful component in the 12om methodology.* Its major strength is that it is well-balanced; yielding the most satisfactory outcome even if this component is not ranked first in terms of feasibility or dimensions of support.

According to the sensitivity analysis, the four main areas that would benefit the most from improvement are:

- Impact on taskwork (which was already relatively high);
- Integration of Perspectives;
- Common Understanding; and
- Collaboration Support.

4.1.2.8 WoG MA Brief Template

WoG MA brief template is 78% satisfactory according to the multi-criteria analysis. This component *has the greatest feasibility* amongst the different components considered here. On the other hand, as revealed by the functional gap analysis, this component is very specific (not flexible) and may not be easily exported to other contexts of use.

According to the sensitivity analysis four equally good areas for improvement are:

- Impact on taskwork;
- OPP support;

- Integration of perspectives;
- Common understanding; and
- Collaboration support.

4.2 Final Remarks

Based on the present results and recommendations, the 12om methodology has completed its evolution in terms of its main components, and with some refinements essentially in terms of training procedure and use simplification, it could indeed be sufficiently mature to be used to good effect in an operational context. We submit that the key to success would be for integrated WoG planning teams to include a specialist capable of driving the application of the 12om methodology and of mitigating any challenges that should arise. The conclusions drawn regarding analyses reported above were taken into account in the development of the 12om process (see ATT5 report entitled "12om Methodology: Process v1.1). The process was designed to mitigate the key issues identified within this report, notably the ones pertaining to time and effort required to use the tools and support to the three dimensions of collaborative understanding.

The team of participants and observers were highly engaged in the OPP simulation of LOE #2. Overall, the pattern of results suggests that the 12om methodology can have a major impact on interagency planning and can be a great asset to Canada for achieving the Comprehensive Approach that it strives to adopt. The 12om methodology is likely to succeed in achieving its desired impact, and the Canadian Forces have already made an official request to DRDC for the right to use this toolkit in its future interagency planning endeavours.

5 References

Benta, M.I. (2005). Studying communication networks with Agna 2.1. *Cognition, Brain, Behavior, 9*(3), 567-574.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen J. D. (2006). The physics of optimal decision making: A formal analysis of performance in two-alternative forced choice tasks. *Psychological Review, 113*, 700-765.

Cannon-Bowers, J.A., & Salas, E. (2001). Reflections on shared cognition. *Journal of Organizational Behavior 22* (2), 195–202.

Endsley, M.R. (1995). Measurement of Situation Awareness in Dynamic Systems. *Human Factors, 37* (1), 65-84.

Gonzalez, C. (2005). Task Workload and Cognitive Abilities in Dynamic Decision Making. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *47*(1), 92-101. doi:10.1518/0018720053653767

Hart, S. G. & Staveland, L. E. (1988) Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.) Human Mental Workload. Amsterdam: North Holland Press.

Hsu, J. S. C., Chang, J. Y. T., Klein, G., & Jiang, J. J. (2011). Exploring the impact of team mental models on information utilization and project performance in system development. *International Journal of Project Management*, *29*(1), 1-12. doi:10.1016/j.ijproman.2009.12.001

Kilian, J., & Siegelmann, H. (1996). The dynamic universality of sigmoidal neural networks. *Information and Computation, 128,* 48-56.

Kuznar, L. A. (2002). Evolutionary applications of risk sensitivity models to socially stratified species: Comparison of sigmoid, concave, and linear. *Evolution and Human Behavior, 23,* 265-280.

Labreuche, C., &, Le Huédé F. (2005). Myriad: a tool suite for MCDA. In International Conference of the Euro Society for Fuzzy Logic and Technology (pp. 204–209), Barcelona, Spain, September 7-9.

Macmillan, J., Paley, M. J., Entin, E. B., & Entin, E. E. (2004). Questionnaires for distributed assessment of team mutual awareness. In N.A. Stanton, A. Hedge, K. brookhuis, E. Salas, H.W. Hendrick (Eds.), *Handbook of human factors and ergonomic methods*. Taylor and Francis.

Matthews, M. D., Beal, S. A, & Pleban, R. J. (2002). Situation Awareness in a Virtual Environment: Description of a Subjective Measure. (Research Report 1786). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Mohammed, S., & Dumville, B.C. (2001). Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behavior 22* (2), 89–106.

Orasanu, J., & Salas, E. (1993). Team decision making in complex environments. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 327-345). Westport, CT, US: Ablex Publishing.

Pew, R.W. (1969). The speed-accuracy operating characteristic. *Acta Psychologica, 30,* 16-26.

Pfautz, J., & E. Roth (2006). Using Cognitive Engineering for System Design and Evaluation: A visualization aid for stability and support operations. *International Journal of Industrial Ergonomics, 36*(5), 389-407.

Potter, S. S., Elm, W., & Tittle, J. (2010). Evaluating the resilience of a human-computer decision-making team: A methodology for decision-centered testing. *Macrocognition,* 1-17.

Powers, E., Stech, F., & Burns, K. (2010). A Behavioral Model of Team Sensemaking. *The International C2 Journal, 4* (1)

Rasmussen, J., A.M. Pejtersen & L.P. Goodstein (1994). *Cognitive Systems Engineering.* John Wiley and Sons, Inc. New-York, NY, USA.

Sharp, M. E., Viswanathan, J., Lanyon, L. J., & Barton, J. J. S. (2012). Sensitivity and bias in decision-making: Evaluating the perception of reward, its probability, and value. *PLoS One, 7,* 1-9.

Wasserman, S., & Faust, K. (1999). *Social Network Analysis: Methods and Applications*. Cambridge University Press, Cambridge, UK

Watts, D. (1999). *Small Worlds*. Princeton University Press, Princeton

Watts, D., & Strogatz, S. H. (1998). Collective Dynamics of Small-World Networks. *Nature, 393*, 440-442.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of Sensemaking. *Organization Science*, *16*(4), 409-421. Wiley-Blackwell. doi:10.1287/orsc.1050.0133

Woods, D. D. & Roth, E. M. (1988). Cognitive Engineering: Human Problem Solving with Tools. *Human Factors, 30* (4), 415-430.

Wickens C. D., & Hollands J. G. (2000). Engineering Psychology and Human Performance (3rd edition). NJ: Prentice Hall.

Appendix A - 12om Methodology Assessment

Methodology assessment

Participant:

Page 1 of 1

Please rate the following items by considering the 12om methodology as a whole.

On a scale ranging from 0 (no support) to 10 (best support), how well does the 12om methodology support the **integration of different perspectives**?

0 1 2 3 4 5 6 7 8 9 10

Integration involves connecting/**putting together several points of view** (or lines of operations) along with their specific perspectives, in the pursuit of a common mission.

On a scale ranging from 0 (no support) to 10 (best support), how well does the 12om methodology support **common understanding**?

0 1 2 3 4 5 6 7 8 9 10

Common understanding is the ability of multiple agents to exploit **common bodies of knowledge, or collective comprehension** for accomplishing a common mission.

On a scale ranging from 0 (no support) to 10 (best support), how well does the 12om methodology support **collaboration**?

0 1 2 3 4 5 6 7 8 9 10

Collaboration is the act of **working with others (coordination, information sharing, deliberation)** to accomplish a common mission.

Appendix B – Component assessment questionnaire

– Cross-impact method shown, but an identical questionnaire was handed out for each component.

Components assessment

Cross-impact method

Participant:

Page 1 of 3

Please rate the **usefulness** of the **cross-impact method**, on the following dimensions using, this scale: 0 (very low usefulness); 5 (moderate usefulness); 10 (very high usefulness)

Usefulness is the ability of something to satisfy needs.

	0 1 2 3 4 5 6 7 8 9 10
Usefulness for supporting the activities and products of the OPP	☐☐☐☐☐☐☐☐☐☐☐
Usefulness for supporting the collaboration process (in general)	☐☐☐☐☐☐☐☐☐☐☐
Usefulness for supporting the integration of different perspectives (in general)	☐☐☐☐☐☐☐☐☐☐☐
Usefulness for supporting the development of common understanding (in general)	☐☐☐☐☐☐☐☐☐☐☐

Please rate the **time and effort required** for applying and training on the **cross-impact method** using this scale: 0 (very little time/effort required); 5 (moderate amount of time/effort required); 10 (a lot of time/effort required)

	0 1 2 3 4 5 6 7 8 9 10
How much **time** (relatively to other tools that may be used in this context) is required for the team to **use** the **cross-impact method**?	☐☐☐☐☐☐☐☐☐☐☐
How much **effort** (relatively to other tools that may be used in this context) is required for the team to **use** the **cross-impact method**?	☐☐☐☐☐☐☐☐☐☐☐
How much **time** (relatively to other tools that may be used in this context) is required for **training** on the **cross-impact method**?	☐☐☐☐☐☐☐☐☐☐☐
How much **effort** (relatively to other tools that may be used in this context) is required for **training** on the **cross-impact method**?	☐☐☐☐☐☐☐☐☐☐☐

What "*other tools*" did you use for comparison in answering the four questions above? List them:

Participant:

Page 2 of 3

Consider the costs and benefits associated with using the cross-impact method. On a scale ranging from 0 (costs significantly outweigh the benefits) to 10 (benefits significantly outweigh the costs), with 5 being a neutral point (costs are equal to benefits), how much do the benefits outweigh the costs?

0 1 2 3 4 5 6 7 8 9 10

Please rate the **usability** of **the cross-impact method** using this scale: 0 (lower usability than similar tools); 5 (equivalent usability than similar tools); 10 (higher usability than similar tools).
Usability is the ease of use of the component, without regard to its usefulness.

What is the **usability** (relatively to other tools that may be used in this context) of the **cross-impact method** component?

0 1 2 3 4 5 6 7 8 9 10

Participant:

Page 3 of 3

Please list **strengths, weaknesses, implementation threats** and **suggested modifications** as well as any other **comments** you would like to share about the **cross-impact method** component.

Strengths:

Weaknesses:

Implementation threats:

Suggested modifications:

Other Comments:

Appendix C – Key mission factors - initial

PARTICIPANT:_____

KEY MISSION FACTORS

From your perspective, what were the 10 most important factors (either favourable or unfavourable) relevant to the achievement of the desired end state of the mission of containing the polio outbreak in Somalia.

1.

2.

3.

4.

5.

6.

7.

8.

9.

10.

Page 1 of 3

Identify 3 most important factors (either favourable or unfavourable) that you think were relevant to the achievement of the desired end state of the mission for the other members of the team.

Team member: _____

1.

2.

3.

Team member: _____

1.

2.

3.

Team member: _____

1.

2.

3.

Team member: _____

1.

2.

3.

Appendix D – Factors linked to mission achievement

– Page 1 shown here, but the actual questionnaire probed participants on 31 factors.

Participant : _____

Factors linked to mission achievement

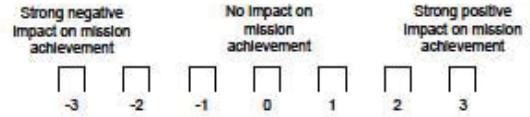
Rate the possible impact of each factor on a 7-point Likert scale.

0 = *No* impact of the factor on mission achievement

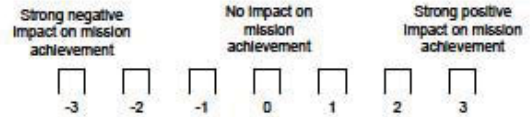
+3 = Strong *positive* impact on mission achievement

-3 = Strong *negative* impact on mission achievement

Factor #1: Stability

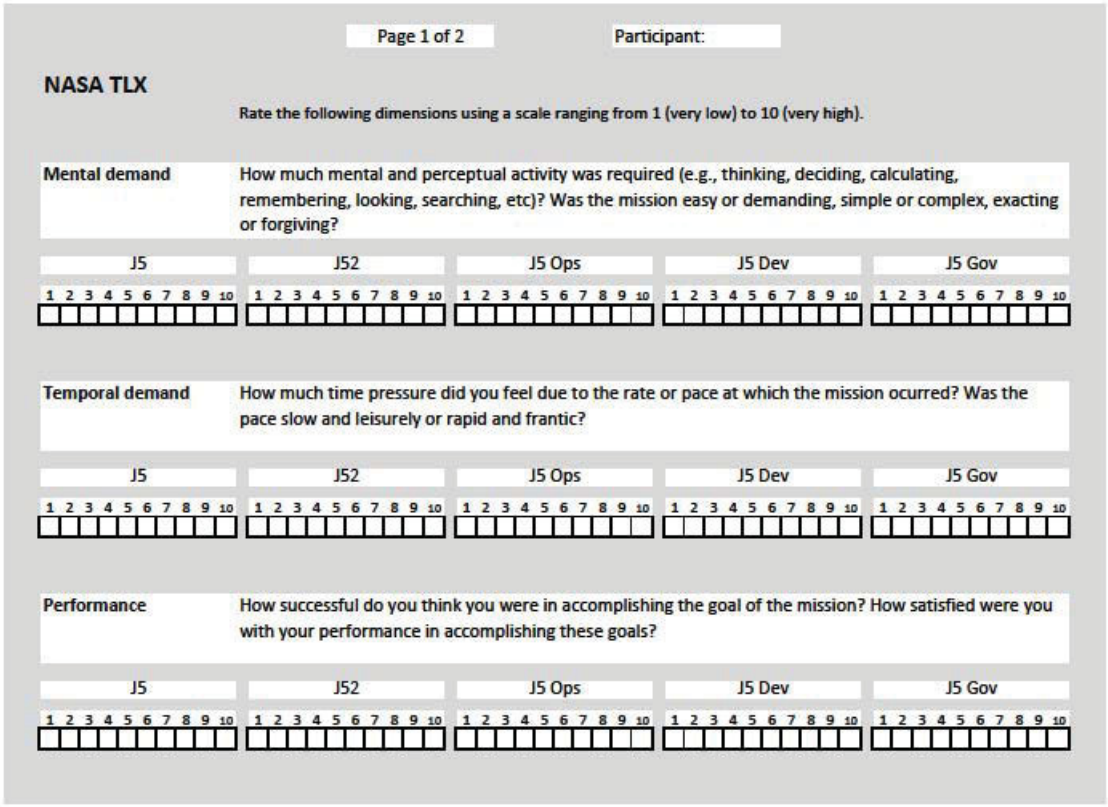


Factor #2: Clan cooperation



Factor #3: FSG capacity to deliver/conduct

Appendix E – NASA-TLX



Page 1 of 2 Participant:

NASA TLX

Rate the following dimensions using a scale ranging from 1 (very low) to 10 (very high).

Mental demand How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc)? Was the mission easy or demanding, simple or complex, exacting or forgiving?

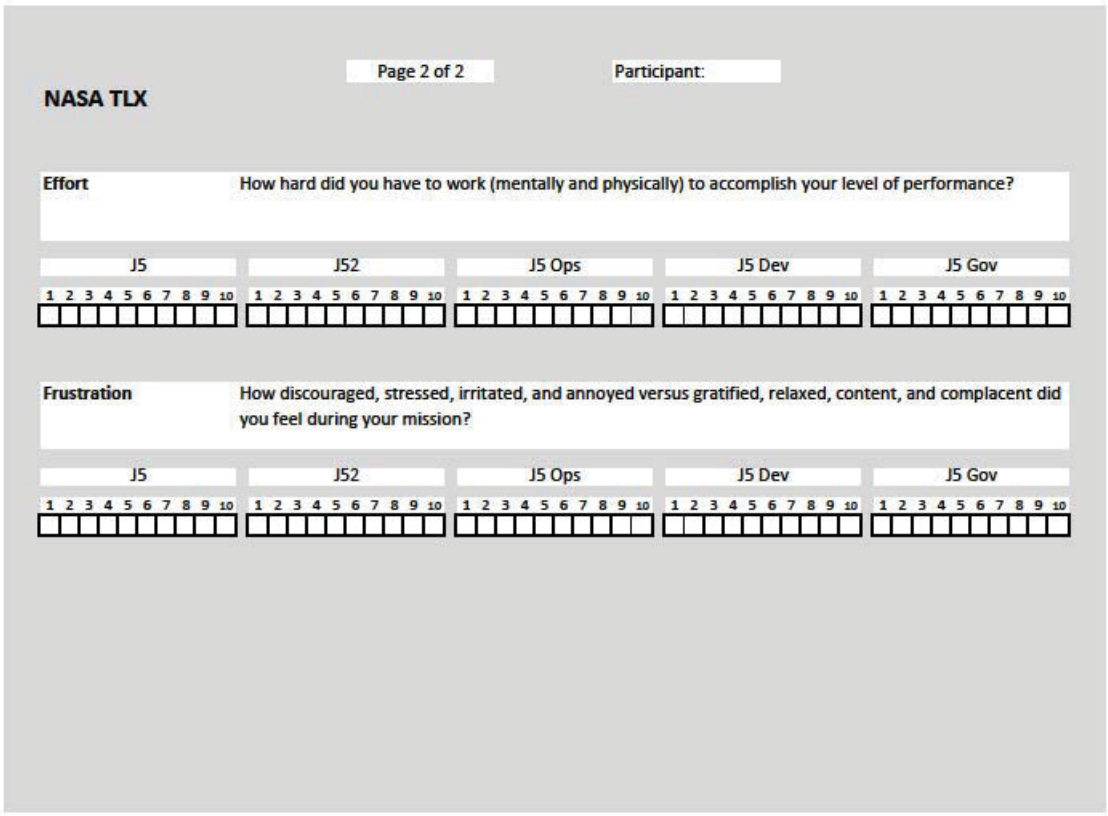
J5	J52	J5 Ops	J5 Dev	J5 Gov
1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10

Temporal demand How much time pressure did you feel due to the rate or pace at which the mission ocurred? Was the pace slow and leisurely or rapid and frantic?

J5	J52	J5 Ops	J5 Dev	J5 Gov
1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10

Performance How successful do you think you were in accomplishing the goal of the mission? How satisfied were you with your performance in accomplishing these goals?

J5	J52	J5 Ops	J5 Dev	J5 Gov
1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10



Page 2 of 2 Participant:

NASA TLX

Effort How hard did you have to work (mentally and physically) to accomplish your level of performance?

J5	J52	J5 Ops	J5 Dev	J5 Gov
1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10

Frustration How discouraged, stressed, irritated, and annoyed versus gratified, relaxed, content, and complacent did you feel during your mission?

J5	J52	J5 Ops	J5 Dev	J5 Gov
1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10	1 2 3 4 5 6 7 8 9 10

Appendix F – TMS

Transactive Memory System

Participant: Page 1 of 1

Please rate your degree of agreement with the following statements on a scale of 1-5 :
(1=strongly disagree, 2=disagree, 3=neutral, 4=agree, 5=strongly agree)

I was comfortable accepting procedural suggestions from other team members.	1 2 3 4 5
I trusted that other members' knowledge about the project was credible.	1 2 3 4 5
I was confident relying on the information that other team members brought to the discussion.	1 2 3 4 5
When other members gave information, I wanted to double-check it for myself.	1 2 3 4 5
I did not have much faith in other members' "expertise".	1 2 3 4 5
Our team worked together in a well-coordinated fashion.	1 2 3 4 5
Our team had very few misunderstandings about what to do.	1 2 3 4 5
Our team needed to backtrack and start over a lot.	1 2 3 4 5
We accomplished the task smoothly and efficiently.	1 2 3 4 5
There was much confusion about how we would accomplish the task.	1 2 3 4 5

Appendix G – MARS

MISSION AWARENESS RATING SCALE (MARS)

1. Please rate your ability to identify mission-critical cues in this mission.

___ very high – able to identify all cues

___ fairly high – could identify most cues

___ somewhat low – many cues hard to identify

___ very low – had substantial problems identifying most cues

2. How difficult – in terms of mental effort required - was it for you to identify or detect mission-critical cues in the mission?

___ very easy – could identify relevant cues with little effort

___ fairly easy – could identify relevant cues, but some effort required

___ somewhat difficult - some effort was required to identify most cues

___ very difficult – substantial effort required to identify relevant cues

3. How well did you understand what was going on during the mission?

___ very well – fully understood the situation as it unfolded

___ fairly well - understood most aspects of the situation

___ somewhat poorly – had difficulty understanding much of the situation

___ very poorly – the situation did not make sense to me

4. How difficult – in terms of mental effort – was it to understand what was going on during the mission?

___ very easy – understood what was going on with little effort

___ fairly easy – understood events with only moderate effort

___ somewhat difficult – hard to comprehend some aspects of the situation

___ very difficult – hard to understand most or all aspects of the situation

5. How difficult – in terms of mental effort – was it to predict what was about to happen during the mission?

___ very easy – little or no effort needed

___ fairly easy – moderate effort required

___ somewhat difficult – many projections required substantial effort

___ very difficult – substantial effort required on most or all projections

6. How well could you predict what was about to occur next in the mission?

___ very well – could predict with accuracy what was about to occur

___ fairly well – could make accurate predictions most of the time

___ somewhat poor – misunderstood the situation much of the time

___ very poor – unable to predict what was about to occur

7. How aware were you of how to best achieve your goals during this mission?

___ very aware – knew how to achieve goals at all times

___ fairly aware – knew most of the time how to achieve mission goals

___ somewhat unaware – was not aware of how to achieve some goals

___ very unaware – generally unaware of how to achieve goals

8. How difficult – in terms of mental effort – was it to decide on how to best achieve mission goals during this mission?

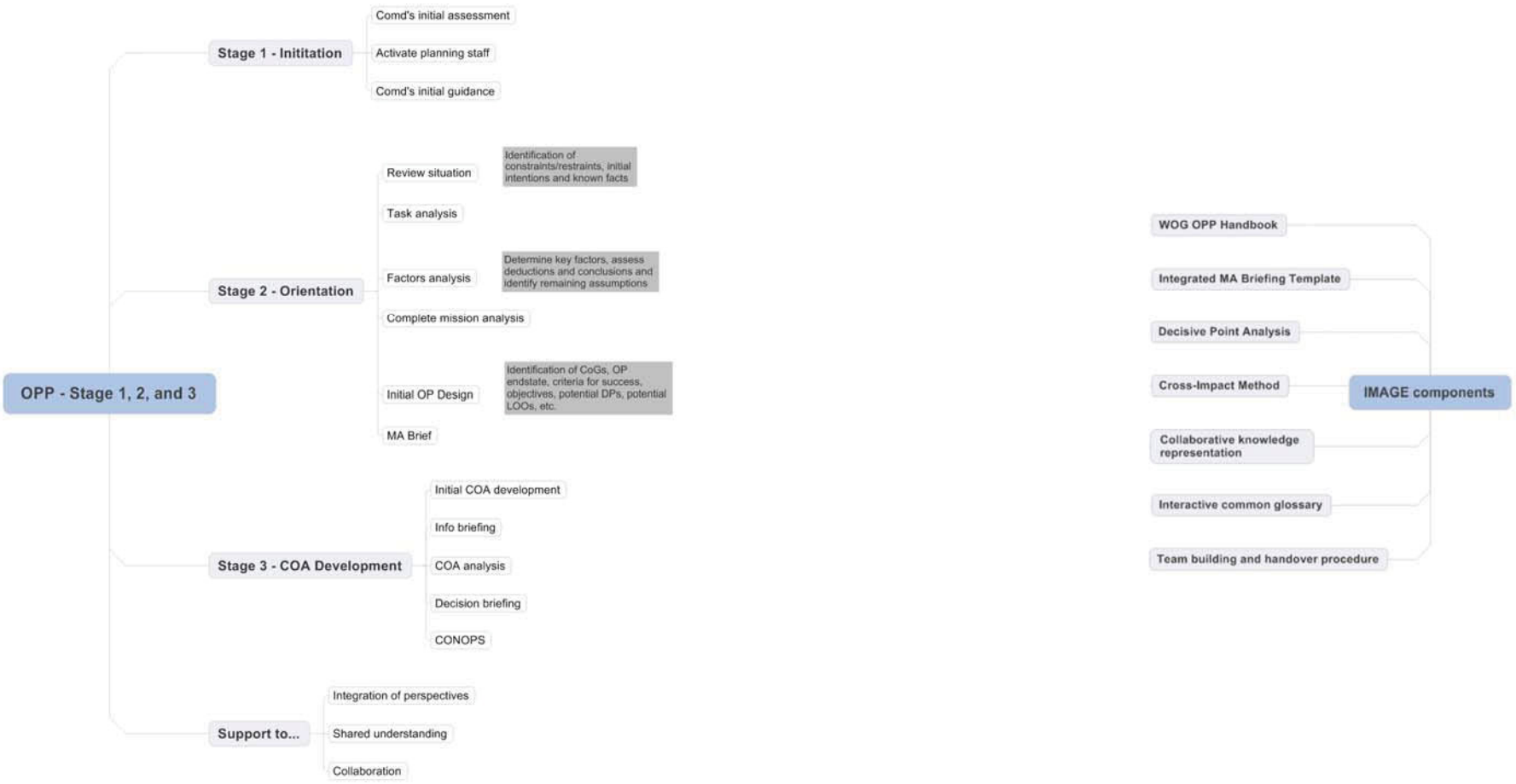
___ very easy – little or no effort needed

___ fairly easy – moderate effort required

___ somewhat difficult – substantial effort needed on some decisions

___ very difficult – most or all decisions required substantial effort

Appendix H – Task to tool mapping



Appendix I – Process evaluation

TEAM: ☐**1** ☐**2**

J5 TEAM PROCESS EVALUATION

Rate the team's process on a 5-point scale from 1 (poor) to 5 (excellent) with 3 being average. Add comments in the space provided

1. How well did the team achieve a common understanding of the situation? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

2. How thoroughly did the team consider higher level direction? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

3. How well did the team identify its own assumptions? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

4. How thoroughly did the team challenge its own assumptions? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

5. How effectively did the team identify constraints and restraints? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

6. How effectively did the team identify key strengths and weaknesses (own and enemy)? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

7. How effectively did the team identify assigned and implied tasks? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

8. How effectively did the team identify objectives? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

9. How effectively did the team describe the desired end state? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

10. How effectively did the team examine force capabilities and groupings? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

11. How effectively did the team consider the command and control structure required? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

12. How effectively did the team assess risk? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

13. How effectively did the team consider the impact of time on the operation? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

14. Was the mission statement complete and explanatory? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

15. Were requests for information (RFI) relevant, clearly stated and the response appropriately incorporated into planning? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

16. Resolution of conflicting information:

- How actively did the team look for conflicting (contradicting) information? | 1 | 2 | 3 | 4 | 5 |
- How easily did the team identify conflicting information? | 1 | 2 | 3 | 4 | 5 |
- How effectively did the team resolve conflicting information? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

17. How easily did the team adapt to changing situations and requirements? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

18. How effectively did the team distribute tasks among its members? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

19. Were all the team members given sufficient opportunity to express their views? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

20. Were these views adequately considered by the rest of the team? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

21. Were all the team members comfortable with their roles and tasks? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

22. How positive was the overall attitude in the team towards:

- The task? | 1 | 2 | 3 | 4 | 5 |
- The team? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

23. Was there effective leadership in the team? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

24. Did the team work effectively together? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

25. How effectively did the team use available time (i.e., an appropriate balance between research, discussion and briefing preparation)? | 1 | 2 | 3 | 4 | 5 |

Comments:_____
_____
_____

26. How effectively did the team adapt the planning process to situational requirements?

1	2	3	4	5

Comments:_____
_____
_____

Appendix J – MA Brief evaluation

EVALUATOR: ☐**COMD** ☐**ROC** ☐**DEV** **TEAM:** ☐**1** ☐**2**

Please fill out the questionnaire below

MISSION ANALYSIS BRIEF EVALUATION CRITERIA

A. Relevance: (by circling one number on each row) **Rate the degree to which the mission analysis brief provides:**

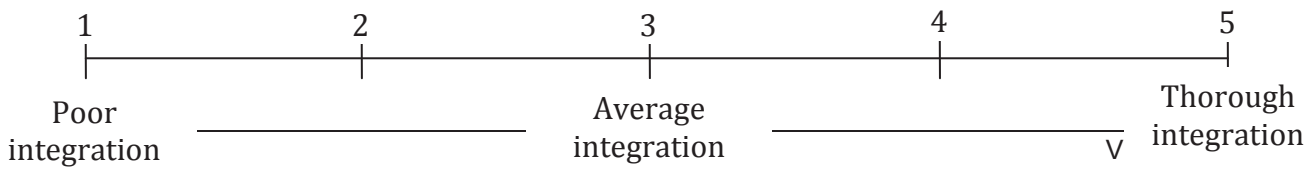
		Does not meet expectations		Meets expectations		Exceeds expectations
A1	The level of detail expected by the Comd	1	2	3	4	5
A2	The level of detail expected by the RoC?	1	2	3	4	5
A3	Quality in the analysis expected by the Comd	1	2	3	4	5
A4	Quality in the analysis expected by the RoC	1	2	3	4	5
A5	Overall, how well does the mission analysis brief meet the Comd's guidance and intent?	1	2	3	4	5
A6	Overall, how well does the mission analysis brief meet the RoC's guidance and intent?	1	2	3	4	5

B. Degree of integration of the three LOOs:

B1. Which LOOs are integrated in the brief (circle one):

i. All three: Security, Development, and Governance
ii. Security and Governance, but not Development
iii. Security and Development, but not Governance
iv. Governance and Development, but not Security
v. None: Each of the three LOOs is included with minimal integration with other LOOs.

B2. If i-iv was chosen in B1, rate the degree of integration by circling one number on the following scale:



C. For each LOO, rate the degree to which the mission analysis brief considered and utilized relevant information, key constraints and restraints, impacts, and provided sufficient level of analysis. *Circle one number in each row*.

C1. Consideration of relevant information/intelligence from the scenario, RFIs, and open source.

		Not considered at all		Considered to some extent		Considered thoroughly
C1.1	Security LOO	1	2	3	4	5
C1.2	Governance LOO	1	2	3	4	5
C1.3	Development LOO	1	2	3	4	5

C2. Consideration of key constraints and restraints.

		Not considered at all		Considered to some extent		Considered thoroughly
C2.1	Security LOO	1	2	3	4	5
C2.2	Governance LOO	1	2	3	4	5
C2.3	Development LOO	1	2	3	4	5

C3. Consideration of the impact of IDP resettlement strategies on

		Not considered at all		Considered to some extent		Considered thoroughly
C3.1	IDPs themselves	1	2	3	4	5
C3.2	Local population	1	2	3	4	5

C4. Analysis of factors and impacts on the Canadian and NATO mission with linkages to considerations for and impacts to tactical and strategic levels

		Not analyzed at all		Analyzed to some extent		Analyzed thoroughly
C4.1	Security LOO	1	2	3	4	5
C4.2	Governance LOO	1	2	3	4	5
C4.3	Development LOO	1	2	3	4	5

D. Sustainability: Rate the extent to which the mission analysis brief gives consideration for transition that retains positive control, sustains results and preserves relationships for each LOO (circle one number in each row):

		Not considered at all		Considered to some extent		Considered thoroughly
D1	Security LOO	1	2	3	4	5
D2	Governance LOO	1	2	3	4	5
D3	Development LOO	1	2	3	4	5

E. OPP manual Mission analysis brief criteria

E1. Mission analysis brief presentation (circle one number in each row):

		Poor		Average		Excellent
E1.1	Clarity of presentation	1	2	3	4	5
E1.2	Coherence of presentation	1	2	3	4	5
E1.3	Clearly stated the aim of the brief	1	2	3	4	5
E1.4	Provided a clear outline of the brief	1	2	3	4	5
E1.5	Use of appropriate visual aids	1	2	3	4	5
E1.6	Appropriate amount of detail	1	2	3	4	5
E1.7	Length	1	2	3	4	5
E1.8	Balance of delivery between the 3 LOOs	1	2	3	4	5

E2. Mission analysis brief content (circle one number in each row)

		Poor		Average		Excellent
E2.1	**Review of situation**					
	a. General					
	i. Security LOO	1	2	3	4	5
	ii. Governance LOO	1	2	3	4	5
	iii. Development LOO	1	2	3	4	5
	b. Review of opposing force situation					
	i. Intent	1	2	3	4	5
	ii. Capabilities	1	2	3	4	5
	iii. Assessment (brief description of likely actions – not COAs)	1	2	3	4	5
	c. Higher Commander's Direction and Guidance					
	(1) Mission for					

	i. Comd	1	2	3	4	5
	ii. RoC	1	2	3	4	5
	(2) Intent for					
	i. Comd	1	2	3	4	5
	ii. RoC	1	2	3	4	5
	(3) Objectives for					
	i. Security LOO	1	2	3	4	5
	ii. Governance LOO	1	2	3	4	5
	iii. Development LOO	1	2	3	4	5
	(4) Transition Conditionsfor					
	i. Security LOO	1	2	3	4	5
	ii. Governance LOO	1	2	3	4	5
	iii. Development LOO	1	2	3	4	5
	(5) Assumptions for					
	i. Security LOO	1	2	3	4	5
	ii. Governance LOO	1	2	3	4	5
	iii. Development LOO	1	2	3	4	5
	(6) Limitations for					
	i. Security LOO	1	2	3	4	5
	ii. Governance LOO	1	2	3	4	5
	iii. Development LOO	1	2	3	4	5
	(7) Assigned Tasks for					
	i. COMD	1	2	3	4	5
	ii. RoC	1	2	3	4	5
	(8) Implied Tasks for					
	i. COMD	1	2	3	4	5
	ii. RoC	1	2	3	4	5
E2.2	**Operations Design**					
	a. Key Factors and Deductions					
	i. Security LOO	1	2	3	4	5
	ii. Governance LOO	1	2	3	4	5
	iii. Development LOO	1	2	3	4	5
	b. Planning Assumptions for					
	i. Security LOO	1	2	3	4	5
	ii. Governance LOO	1	2	3	4	5
	iii. Development LOO	1	2	3	4	5
	c. Key Strengths and Weaknesses					
	i. Blue force	1	2	3	4	5
	ii. Red force	1	2	3	4	5
	iii. Green force	1	2	3	4	5
	iv. White force	1	2	3	4	5
	d. Strategic Centers of gravity for					
	i. Blue force	1	2	3	4	5
	ii. Red force	1	2	3	4	5

	iii. Green force	1	2	3	4	5
	iv. White force	1	2	3	4	5
	e. Operational Centers of gravity for					
	i. Blue force	1	2	3	4	5
	ii. Red force	1	2	3	4	5
	iii. Green force	1	2	3	4	5
	iv. White force	1	2	3	4	5
	f. Decisive Points					
	i. Security LOO	1	2	3	4	5
	ii. Governance LOO	1	2	3	4	5
	iii. Development LOO	1	2	3	4	5
	g. Objectives integrated for					
	i. Security LOO	1	2	3	4	5
	ii. Governance LOO	1	2	3	4	5
	iii. Development LOO	1	2	3	4	5
	h. End State and Criteria for Success integrated for					
	i. Security LOO	1	2	3	4	5
	ii. Governance LOO	1	2	3	4	5
	iii. Development LOO	1	2	3	4	5
E2.3	**Proposed mission statement for**					
	i. Comd	1	2	3	4	5
	ii. RoC	1	2	3	4	5
E2.4	**Initial Force Estimate**					
	i. Estimate of Forces Required	1	2	3	4	5
E2.5	**Proposed Planning Guidance**					
	a. Proposed Initial Intent					
	i. COMD	1	2	3	4	5
	ii. RoC	1	2	3	4	5
	b. Direction to Planning Staff (Staff Planning Directive) integrated for					
	i. Security LOO	1	2	3	4	5
	ii. Governance LOO	1	2	3	4	5
	iii. Development LOO	1	2	3	4	5
	c. Direction to Subordinate Commands (Wng O) integrated for					
	i. Security LOO	1	2	3	4	5
	ii. Governance LOO	1	2	3	4	5
	iii. Development LOO	1	2	3	4	5
E2.6	**Conclusion**	1	2	3	4	5

F. RATE THE OVERALL QUALITY OF THE MISSION ANALYSIS BRIEF (circle one number)

1 2 3 4 5

Poor Average Excellent

COA EVALUATION CRITERIA

G. COA evaluation criteria (circle one number in each row)

		Poor		Average		Excellent
G1	Time required to make decision	1	2	3	4	5
G2	Shared vision between the planning team in designing the COAs	1	2	3	4	5
G3	Perception of credibility by:					
	i. Comd	1	2	3	4	5
	ii. RoC	1	2	3	4	5
G4	Degree of originality of the COAs in integrating the constraints and opportunities of each LOO towards an optimum workable COA	1	2	3	4	5
G5	Degree of system coordination across the 3 LOOs	1	2	3	4	5
G6	Degree of levering earlier experience	1	2	3	4	5
G7	Feasibility of the proposed COAs for each LOO					
	i. Security LOO	1	2	3	4	5
	ii. Governance LOO	1	2	3	4	5
	iii. Development LOO	1	2	3	4	5
G8	Feasibility of the proposed COAs as joint 3 LOOs	1	2	3	4	5
G9	Variability among the proposed COAs	1	2	3	4	5
G10	Degree to which conflicting objectives from each LOO were incorporated to form feasible joint COAs:					
	i. Security LOO	1	2	3	4	5
	ii. Governance LOO	1	2	3	4	5
	iii. Development LOO	1	2	3	4	5
G11	Degree to which each COA achieves the mission objectives for both COMD and RoC					
	COA 1: Achieves objectives for the COMD	1	2	3	4	5
	COA 1: Achieves objectives for the RoC	1	2	3	4	5
	COA 2: Achieves objectives for the COMD	1	2	3	4	5
	COA 2: Achieves objectives for the RoC	1	2	3	4	5
	COA 3: Achieves objectives for the COMD	1	2	3	4	5
	COA 3: Achieves objectives for the RoC	1	2	3	4	5
G12	Conducted Factors analysis	1	2	3	4	5
G13	Developed several COAs	1	2	3	4	5

G14	Developed COA comparison criteria for					
	i. Security LOO	1	2	3	4	5
	ii. Governance LOO	1	2	3	4	5
	iii. Development LOO	1	2	3	4	5
G15	Refined operations design	1	2	3	4	5
G16	Identified decision points	1	2	3	4	5

Appendix K – SME ratings to component assessment by component

	Sharing conceptual diagrams		Querying conceptual diagrams	Creating views using filters	OPP Handbook	Creation of conceptual diagrams collab.	Creating conceptual diagrams individual	WoG MA Briefing template	Cross-impact method	OP design tool	Common Glossary	Team building	OP Design process	Creation of the common vocabulary
Usefulness for supporting the activities and products of the OPP	7	na	na	8	5	8	7	8	7	7	6	8	5	
Usefulness for supporting the collaboration process (in general)	9	na	na	8	9	7	6	7	4	6	8	8	6	
Usefulness for supporting the integration of different perspectives (in general)	9	na	na	7	10	8	7	8	5	7	6	8	6	
Usefulness for supporting the development of common understanding (in general)	10	na	na	9	10	6	7	8	5	8	6	8	6	
How much time is required for the team to use this component	2	na	na	4	2	4	3	6	4	3	3	5	2	
How much effort is required for the team to use this component	4	na	na	4	2	4	4	6	4	3	3	6	2	
How much time is required for training on this component	3	na	na	2	3	4	2	8	3	2	0	0	1	
How much effort is required for training on this component	3	na	na	3	4	4	1	8	3	3	0	7	1	